

STIC-Biotech/ChemLib

142525

From: Priebe, Scott
Sent: Thursday, January 13, 2005 9:24 AM
To: STIC-Biotech/ChemLib
Cc: Whiteman, Brian
Subject: FW: sequence search - RUSH

Importance: High

Please RUSH the requested search.

-----Original Message-----

From: Whiteman, Brian
Sent: Thursday, January 13, 2005 9:10 AM
To: STIC-Biotech/ChemLib; Priebe, Scott
Subject: sequence search

10/057,136
EFD 2/24/97
Inventor: Schlom et al.

Please perform an oligonucleotide search for SEQ ID NOs: 2 and 4-12 against public databases, us patent, and us patent application databases.

Thank you,

Brian Whiteman
Remsen, 2D14
mail box 2C18
Patent Examiner - Art Unit 1635
United States Patent and Trademark Office
(571) 272-0764

STAFF USE ONLY

Searcher: _____
Searcher Phone: 2- _____
Date Searcher Picked up: _____
Date Completed: _____
Searcher Prep/Rev. Time: _____
Online Time: _____

Type of Search

NA Sequence: # _____
AA Sequence: # _____
Structure: # _____
Bibliographic: _____
Litigation: _____
Patent Family: _____
Other: _____

Vendors and cost where applicable

STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other(Specify): _____

This Page Blank (uspto)

Db 1 GGTAGTACAGCGCCACCGCACATGCGCTCAGAGCGCTCCGGATACGAGACCGGCGCCT 60

RESULT 2
US-10-406-317-41
Sequence 41, Application US/10406317
Publication No. US20040019195A1
GENERAL INFORMATION:
APPLICANT: Schlom, Jeffrey;
APPLICANT: Hodge, James;
APPLICANT: Panicali, Dennis
TITLE OF INVENTION: A recombinant vector expressing multiple constimulatory
FILE REFERENCE: 38163-0189
CURRENT APPLICATION NUMBER: US/10/406,317
CURRENT FILING DATE: 2003-04-04
PRIOR APPLICATION NUMBER: US/09/856,988
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: PCT/US99/26866
PRIOR FILING DATE: 1999-11-12
PRIOR APPLICATION NUMBER: 60/111,582
PRIOR FILING DATE: 1998-12-09
NUMBER OF SEQ ID NOS: 41
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 41
LENGTH: 2297
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: VECTOR
US-10-406-317-41

Query Match
Best Local Similarity 100.0%; Score 60; DB 16; Length 2297;
Matches 60; Conservative 100.0%; Pred. No. 4.5e-13;
Matches 60; Indels 0; Gaps 0;

Qy 1 GGTAGTACAGCGCCACCGCACATGCGCTCAGAGCGCTCCGGATACGAGACCGGCGCCT 60
Db 886 GGTAGTACAGCGCCACCGCACATGCGCTCAGAGCGCTCCGGATACGAGACCGGCGCCT 945

RESULT 3
US-10-057-136-9
Sequence 9, Application US/10057136
Publication No. US20030021770A1
GENERAL INFORMATION:
APPLICANT: SCHLOM, JEFFREY
APPLICANT: KANTOR, JUDITH
APPLICANT: KUBE, DONALD
APPLICANT: PANICALI, DENNIS
APPLICANT: GRITZ, LINDA
TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
FILE REFERENCE: 700953/47113C
CURRENT APPLICATION NUMBER: US/10/057,136
CURRENT FILING DATE: 2002-01-25
PRIOR APPLICATION NUMBER: 09/366,670
PRIOR FILING DATE: 1999-08-03
PRIOR APPLICATION NUMBER: PCT/US98/03693
PRIOR FILING DATE: 1998-02-24
PRIOR APPLICATION NUMBER: 60/038,253
PRIOR FILING DATE: 1997-02-24
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 9
LENGTH: 60
TYPE: DNA
ORGANISM: Homo sapiens
US-10-057-136-9

Query Match
Best Local Similarity 63.7%; Score 38.2; DB 14; Length 60;
Matches 78.0%; Pred. No. 8.6e-05;

Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 1 GGTAGTACAGCGCCACCGCACATGCGCTCAGAGCGCTCCGGATACGAGACCGGCGCCT 59
Db 1 GGAAGTACCGCTCTCACTGTCACACGCGGCTCAAGCGCGCCAGACATCTCGCC 59

RESULT 4
US-10-057-136-4
Sequence 4, Application US/10057136
Publication No. US20030021770A1
GENERAL INFORMATION:
APPLICANT: SCHLOM, JEFFREY
APPLICANT: KANTOR, JUDITH
APPLICANT: KUBE, DONALD
APPLICANT: PANICALI, DENNIS
APPLICANT: GRITZ, LINDA
TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
FILE REFERENCE: 700953/47113C
CURRENT APPLICATION NUMBER: US/10/057,136
CURRENT FILING DATE: 2002-01-25
PRIOR APPLICATION NUMBER: 09/366,670
PRIOR FILING DATE: 1999-08-03
PRIOR APPLICATION NUMBER: PCT/US98/03693
PRIOR FILING DATE: 1998-02-24
PRIOR APPLICATION NUMBER: 60/038,253
PRIOR FILING DATE: 1997-02-24
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 60
TYPE: DNA
ORGANISM: Homo sapiens
US-10-057-136-4

Query Match
Best Local Similarity 62.7%; Score 37.6; DB 14; Length 60;
Matches 46; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 1 GGTAGTACAGCGCCACCGCACATGCGCTCAGAGCGCTCCGGATACGAGACCGGCGCCT 60
Db 1 GGCAGTACTGCACACCGGACATGCGCTGAATCATGACACTGATCAAGACTTGCACCT 60

RESULT 5
US-10-057-136-8
Sequence 8, Application US/10057136
Publication No. US20030021770A1
GENERAL INFORMATION:
APPLICANT: SCHLOM, JEFFREY
APPLICANT: KANTOR, JUDITH
APPLICANT: KUBE, DONALD
APPLICANT: PANICALI, DENNIS
APPLICANT: GRITZ, LINDA
TITLE OF INVENTION: TUMOR-ASSOCIATED ANTIGEN
FILE REFERENCE: 700953/47113C
CURRENT APPLICATION NUMBER: US/10/057,136
CURRENT FILING DATE: 2002-01-25
PRIOR APPLICATION NUMBER: 09/366,670
PRIOR FILING DATE: 1999-08-03
PRIOR APPLICATION NUMBER: PCT/US98/03693
PRIOR FILING DATE: 1998-02-24
PRIOR APPLICATION NUMBER: 60/038,253
PRIOR FILING DATE: 1997-02-24
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 8
LENGTH: 60
TYPE: DNA
ORGANISM: Homo sapiens
US-10-057-136-8

Query Match 61.0%; Score 36.6; DB 14; Length 60;
Best Local Similarity 76.3%; Pred. No. 0.00035;
Matches 45; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 1 GGTAAGACAGCCGCCGACATGCGCTCAGAGCGCTCCGATACGAGACCGGCGCC 59
Db 1 GGCAGACCGCACCGCCGACACGGGTCACAAAGCGCGCACACTCGACTCGCGCC 59

RESULT 6

US-10-057-136-7
Sequence 7, Application US/10057136
Publication No. US20030021770A1

GENERAL INFORMATION:

APPLICANT: SCHLOM, JEFFREY
APPLICANT: KANTOR, JUDITH

APPLICANT: KUPFE, DONALD
APPLICANT: PANICALI, DENNIS

APPLICANT: GRITZ, LINDA

TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1

FILE REFERENCE: 700953/47113C

CURRENT APPLICATION NUMBER: US/10/057,136

PRIOR FILING DATE: 2002-01-25

PRIOR APPLICATION NUMBER: 09/366,670

PRIOR FILING DATE: 1999-08-03

PRIOR APPLICATION NUMBER: PCT/US98/03693

PRIOR FILING DATE: 1998-02-24

PRIOR APPLICATION NUMBER: 60/038,253

PRIOR FILING DATE: 1997-02-24

NUMBER OF SEQ ID NOS: 20

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 7

LENGTH: 60

TYPE: DNA

ORGANISM: Homo sapiens

US-10-057-136-7

Query Match 54.7%; Score 32.8; DB 14; Length 60;
Best Local Similarity 71.7%; Pred. No. 0.0094;
Matches 43; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 1 GGTAAGACAGCCGCCGACATGCGCTCAGAGCGCTCCGATACGAGACCGGCGCT 60
Db 1 GGTTCAGCCGCCCGCTCTCTCAGGTGATCCGCCGATACAGACCGGCGCT 60

RESULT 7

US-10-057-136-6

Sequence 6, Application US/10057136

Publication No. US20030021770A1

GENERAL INFORMATION:

APPLICANT: SCHLOM, JEFFREY
APPLICANT: KANTOR, JUDITH

APPLICANT: KUPFE, DONALD
APPLICANT: PANICALI, DENNIS

APPLICANT: GRITZ, LINDA

TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1

FILE REFERENCE: 700953/47113C

CURRENT APPLICATION NUMBER: US/10/057,136

PRIOR FILING DATE: 2002-01-25

PRIOR APPLICATION NUMBER: 09/366,670

PRIOR FILING DATE: 1999-08-03

PRIOR APPLICATION NUMBER: PCT/US98/03693

PRIOR FILING DATE: 1998-02-24

PRIOR APPLICATION NUMBER: 60/038,253

PRIOR FILING DATE: 1997-02-24

NUMBER OF SEQ ID NOS: 20

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 6

TYPE: DNA
ORGANISM: Homo sapiens
US-10-057-136-6

Query Match 48.3%; Score 29; DB 14; Length 60;
Best Local Similarity 71.7%; Pred. No. 0.26;
Matches 38; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 7 ACAGCGCCACCGGACATGCGCTCAGAGCGCTCCGATACGAGACCGGCGCC 59
Db 7 ACAGCTCTCCGCTCAGTGGGTACTTCTGCTCCAGATATCGCCAGCTCC 59

RESULT 8

US-10-296-734-1165

Sequence 1165, Application US/10296734

Publication No. US20040054137A1

GENERAL INFORMATION:

APPLICANT: Thompson, Scott A
APPLICANT: Ramshaw, Ian A

TITLE OF INVENTION: Synthetic molecules and uses therefor

FILE REFERENCE: Savine

CURRENT APPLICATION NUMBER: US/10/296,734

PRIOR FILING DATE: 2003-08-04

PRIOR APPLICATION NUMBER: AU P07761/00

PRIOR FILING DATE: 2000-05-26

NUMBER OF SEQ ID NOS: 1507

SOFTWARE: PatentIn version 3.2

SEQ ID NO 1165

LENGTH: 72

TYPE: DNA

ORGANISM: Artificial

FEATURE: OTHER INFORMATION: MUC1F segment 8

FEATURE: NAME/KEY: CDS

LOCATION: (1)..(72)

US-10-296-734-1165

Query Match 48.0%; Score 28.8; DB 16; Length 72;
Best Local Similarity 69.6%; Pred. No. 0.3;
Matches 39; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 1 GGTAAGACAGCCGCCGACATGCGCTCAGAGCGCTCCGATACGAGACCGGCGC 56
Db 16 GGAAGACACACCCCTCCGCTCAGATGAGACAGCGCTCCGATACAAAGCGC 71

RESULT 9

US-10-296-734-1209

Sequence 1209, Application US/10296734

Publication No. US20040054137A1

GENERAL INFORMATION:

APPLICANT: Thompson, Scott A
APPLICANT: Ramshaw, Ian A

TITLE OF INVENTION: Synthetic molecules and uses therefor

FILE REFERENCE: Savine

CURRENT APPLICATION NUMBER: US/10/296,734

PRIOR FILING DATE: 2003-08-04

PRIOR APPLICATION NUMBER: AU P07761/00

PRIOR FILING DATE: 2000-05-26

NUMBER OF SEQ ID NOS: 1507

SOFTWARE: PatentIn version 3.2

SEQ ID NO 1209

LENGTH: 16638

TYPE: DNA

ORGANISM: Artificial

FEATURE: OTHER INFORMATION: Differentiation svaine

FEATURE: NAME/KEY: CDS

LOCATION: (1)..(16638)

US-10-296-734-1209

Query Match 48.0%; Score 28.8; DB 16; Length 1638;
 Best Local Similarity 69.6%; Pred. No. 0.24;
 Matches 39; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 1 GGTAGTACAGGCGCCGACATGGGCTCAGCAGCGCTCCGATACGAGACCGCG 56
 Db 4246 GGAAGCACACCTCTCCGCTCAGATGTGACAAAGCGCTCCGATACAAAGCCGC 4301

RESULT 10
 ; Sequence 2, Application US/10057136
 ; Publication No. US20030021770A1
 ; GENERAL INFORMATION:
 ; APPLICANT: SCHIOW, JEFFREY
 ; APPLICANT: KANTOR, JUDITH
 ; APPLICANT: KUFE, DONALD
 ; APPLICANT: PANICALI, DENNIS
 ; APPLICANT: GRITZ, LINDA
 ; TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
 ; FILE REFERENCE: 700953/47113C
 ; CURRENT APPLICATION NUMBER: US/10/057,136
 ; PRIOR FILING DATE: 2002-01-25
 ; PRIOR APPLICATION NUMBER: 09/366,670
 ; PRIOR FILING DATE: 1999-08-03
 ; PRIOR APPLICATION NUMBER: PCT/US98/03693
 ; PRIOR FILING DATE: 1998-02-24
 ; PRIOR APPLICATION NUMBER: 60/038,253
 ; PRIOR FILING DATE: 1997-02-24
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 60
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-10-057-136-2

Query Match 45.7%; Score 27.4; DB 14; Length 60;
 Best Local Similarity 69.8%; Pred. No. 1;
 Matches 37; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 7 ACAGCGCACCCGACATGGGCTCCGATACGAGACCGCGCC 59
 Db 7 ACCGCCCCCAGCCACGGTGTCACTCGCCCCGACACGAGCGCGCC 59

RESULT 11
 ; US-10-447-839A-75/c
 ; Sequence 75, Application US/10447839A
 ; Publication No. US20040018181A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kufe, Donald W.
 ; APPLICANT: Kharbada, Surender
 ; APPLICANT: Weisman, Steven D.
 ; TITLE OF INVENTION: MUC1 INTERFERENCE RNA COMPOSITIONS AND METHODS DERIVED THEREFROM
 ; FILE REFERENCE: 1000.05.009
 ; CURRENT APPLICATION NUMBER: US/10/447,839A
 ; PRIOR FILING DATE: 2003-05-29
 ; PRIOR APPLICATION NUMBER: 10/293,391
 ; PRIOR FILING DATE: 2002-11-13
 ; PRIOR APPLICATION NUMBER: 09/951,938
 ; PRIOR FILING DATE: 2001-09-11
 ; PRIOR APPLICATION NUMBER: 60/231,841
 ; PRIOR FILING DATE: 2000-09-11
 ; NUMBER OF SEQ ID NOS: 109
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 75
 ; LENGTH: 1424
 ; TYPE: RNA
 ; ORGANISM: ARTIFICIAL
 ; FEATURE:

OTHER INFORMATION: Synthesized Sequence
 ; US-10-447-839A-75

Query Match 45.7%; Score 27.4; DB 16; Length 1424;
 Best Local Similarity 69.8%; Pred. No. 0.9;
 Matches 37; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 7 ACAGCGCACCCGACATGGGCTCAGCAGCGCTCCGATACGAGACCGCGCC 59
 Db 1035 ACCGCCCCCAGCCACGGTGTCACTCGCCCCGACACGAGCGCGCC 983

RESULT 12
 ; US-10-447-839A-20
 ; Sequence 20, Application US/10447839A
 ; Publication No. US20040018181A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kufe, Donald W.
 ; APPLICANT: Kharbada, Surender
 ; APPLICANT: Weisman, Steven D.
 ; TITLE OF INVENTION: MUC1 INTERFERENCE RNA COMPOSITIONS AND METHODS DERIVED THEREFROM
 ; FILE REFERENCE: 1000.05.009
 ; CURRENT APPLICATION NUMBER: US/10/447,839A
 ; PRIOR FILING DATE: 2003-05-29
 ; PRIOR APPLICATION NUMBER: 10/293,391
 ; PRIOR FILING DATE: 2002-11-13
 ; PRIOR APPLICATION NUMBER: 09/951,938
 ; PRIOR FILING DATE: 2001-09-11
 ; PRIOR APPLICATION NUMBER: 60/231,841
 ; PRIOR FILING DATE: 2000-09-11
 ; NUMBER OF SEQ ID NOS: 109
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 20
 ; LENGTH: 1428
 ; TYPE: RNA
 ; ORGANISM: RNA
 ; US-10-447-839A-20

Query Match 45.7%; Score 27.4; DB 16; Length 1428;
 Best Local Similarity 67.9%; Pred. No. 0.9;
 Matches 36; Conservative 1; Mismatches 16; Indels 0; Gaps 0;

Qy 7 ACAGCGCACCCGACATGGGCTCCGATACGAGACCGCGCC 59
 Db 391 ACCGCCCCCAGCCACGGTGTCACTCGCCCCGACACGAGCGCGCC 443

RESULT 13
 ; US-10-057-136-19
 ; Sequence 19, Application US/10057136
 ; Publication No. US20030021770A1
 ; GENERAL INFORMATION:
 ; APPLICANT: SCHIOW, JEFFREY
 ; APPLICANT: KANTOR, JUDITH
 ; APPLICANT: KUFE, DONALD
 ; APPLICANT: PANICALI, DENNIS
 ; APPLICANT: GRITZ, LINDA
 ; TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
 ; FILE REFERENCE: 700953/47113C
 ; CURRENT APPLICATION NUMBER: US/10/057,136
 ; PRIOR FILING DATE: 2002-01-25
 ; PRIOR APPLICATION NUMBER: 09/366,670
 ; PRIOR FILING DATE: 1999-08-03
 ; PRIOR APPLICATION NUMBER: PCT/US98/03693
 ; PRIOR FILING DATE: 1998-02-24
 ; PRIOR APPLICATION NUMBER: 60/038,253
 ; PRIOR FILING DATE: 1997-02-24
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 19
 ; LENGTH: 1527
 ; TYPE: DNA

ORGANISM: Homo sapiens
US-10-057-136-19

Query Match 45.7%; Score 27.4; DB 14; Length 1527;
Best Local Similarity 69.8%; Pred. No. 0.9;
Matches 37; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 7 ACAGCGCCACCCGACATGGCGTCCGAGCGCTCCGATACGAGACCGGCGCC 59
DB 232 ACCGCGCCCGACGCCACGATGTCACCTCGGCCCCGGACACCGCGCGCC 284

RESULT 14
US-09-864-864-280

Sequence 280, Application US/09864864
Patent No. US20020102679A1
GENERAL INFORMATION:

APPLICANT: Xu, Jiangchun
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan L.
APPLICANT: Dillon, David C.
APPLICANT: Secrist, Heather
APPLICANT: Lodes, Michael J.
APPLICANT: Algate, Paul A.
APPLICANT: Fling, Steve P.
APPLICANT: Mannion, Jane
APPLICANT: Benson, Darin R.
APPLICANT: Carter, Darick
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
FILE REFERENCE: 210121.523
CURRENT APPLICATION NUMBER: US/09/864,864
CURRENT FILING DATE: 2001-05-23
NUMBER OF SEQ ID NOS: 341
SOFTWARE: Corixa Invention Disclosure Database
SEQ ID NO 280
LENGTH: 1721
TYPE: DNA
ORGANISM: Homo sapiens
US-09-864-864-280

Query Match 45.7%; Score 27.4; DB 9; Length 1721;
Best Local Similarity 69.8%; Pred. No. 0.89;
Matches 37; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 7 ACAGCGCCACCCGACATGGCGTCCGAGCGCTCCGATACGAGACCGGCGCC 59
DB 568 ACCGCGCCCGACGCCACGATGTCACCTCGGCCCCGGACACCGCGCGCC 620

RESULT 15
US-09-967-768A-224

Sequence 224, Application US/09967768A
Patent No. US20020150877A1
GENERAL INFORMATION:

APPLICANT: Augustus, Meena
TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
FILE REFERENCE: 689290-72
CURRENT APPLICATION NUMBER: US/09/967,768A
CURRENT FILING DATE: 2001-09-28
PRIOR APPLICATION NUMBER: US/60/236,109
PRIOR FILING DATE: 2000-09-28
PRIOR APPLICATION NUMBER: US/60/236,034
PRIOR FILING DATE: 2000-09-28
PRIOR APPLICATION NUMBER: US/60/236,111
PRIOR FILING DATE: 2000-09-28
NUMBER OF SEQ ID NOS: 325
SOFTWARE: PatentIn version 3.0
SEQ ID NO 224
LENGTH: 1721
TYPE: DNA
ORGANISM: Homo sapiens

US-09-967-768A-224

Query Match 45.7%; Score 27.4; DB 9; Length 1721;
Best Local Similarity 69.8%; Pred. No. 0.89;
Matches 37; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 7 ACAGCGCCACCCGACATGGCGTCCGAGCGCTCCGATACGAGACCGGCGCC 59
DB 568 ACCGCGCCCGACGCCACGATGTCACCTCGGCCCCGGACACCGCGCGCC 620

Search completed: January 16, 2005, 09:30:31
Job time : 183.4 secs

This Page Blank (uspto)

Query Match	45.7%;	Score 27.4;	DB 4;	Length 8186;
Best Local Similarity	69.8%;	Pred. No. 1.4;		

```

RESULT 6
US-10-029-517-18
; Sequence 18, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Susan J. Myers
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RRS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517

```

;/ CURRENT FILING DATE: 2001-12-20
;/ NUMBER OF SEQ ID NOS: 107
;/ SEQ ID NO 18
;/ LENGTH: 572
;/ TYPE: DNA
;/ ORGANISM: Homo sapiens
;/ FEATURE:
;/ NAME/KEY: CDS
;/ LOCATION: (67)...(572)
US-10-029-517-18

Query Match 43.0%; Score 25.8; DB 4; Length 572;
Best Local Similarity 67.9%; Pred. No. 3.9;
Matches 36; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 7 ACAGCCGCCACCCGACATGCGCTCAGAGCGCTCCGGATACGAGACCGGCC 59
Db 484 ACCGCCCCCAAGCCAGCGGTGTACCTCGGCCCGGACACGAGCGGCC 536

RESULT 7
US-10-029-517-102
;/ Sequence 102, Application US/10029517
;/ Patent No. 6716627
;/ GENERAL INFORMATION:
;/ APPLICANT: Kenneth W. Dobie
;/ APPLICANT: Susan J. Myers
;/ TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
;/ FILE REFERENCE: RTS-0352
;/ CURRENT APPLICATION NUMBER: US/10/029,517
;/ CURRENT FILING DATE: 2001-12-20
;/ NUMBER OF SEQ ID NOS: 107
;/ SEQ ID NO 102
;/ LENGTH: 3343
;/ TYPE: DNA
;/ ORGANISM: Homo sapiens
;/ FEATURE:
US-10-029-517-102

Query Match 43.0%; Score 25.8; DB 4; Length 3343;
Best Local Similarity 67.9%; Pred. No. 4.5;
Matches 36; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 7 ACAGCCGCCACCCGACATGCGCTCAGAGCGCTCCGGATACGAGACCGGCC 59
Db 1674 ACCGCCCCCAAGCCAGCGGTGTACCTCGGCCCGGACACGAGCGGCC 1726

RESULT 8
US-10-029-517-16
;/ Sequence 16, Application US/10029517
;/ Patent No. 6716627
;/ GENERAL INFORMATION:
;/ APPLICANT: Kenneth W. Dobie
;/ APPLICANT: Susan J. Myers
;/ TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
;/ FILE REFERENCE: RTS-0352
;/ CURRENT APPLICATION NUMBER: US/10/029,517
;/ CURRENT FILING DATE: 2001-12-20
;/ NUMBER OF SEQ ID NOS: 107
;/ SEQ ID NO 16
;/ LENGTH: 981
;/ TYPE: DNA
;/ ORGANISM: Homo sapiens
;/ FEATURE:
;/ NAME/KEY: exon:exon junction
;/ LOCATION: (464)...(465)
;/ OTHER INFORMATION: exon 3b:exon 4
US-10-029-517-16

Query Match 40.7%; Score 24.4; DB 4; Length 981;
Best Local Similarity 68.0%; Pred. No. 12;
Matches 34; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 7 ACAGCCGCCACCCGACATGCGCTCAGAGCGCTCCGGATACGAGACCGGCC 56
Db 27 ACCGCCCCCAAGCCAGCGGTGTACCTCGGCCCGGACACGAGCGGCC 76

RESULT 9
US-09-475-947A-246
;/ Sequence 246, Application US/09475947A
;/ Patent No. 6472154
;/ GENERAL INFORMATION:
;/ APPLICANT: Garner, Harold R.
;/ APPLICANT: Wren, Jonathan D.
;/ APPLICANT: Minna, John D.
;/ TITLE OF INVENTION: Polymorphic Repeats in Human Genes
;/ FILE REFERENCE: UTS0667
;/ CURRENT APPLICATION NUMBER: US/09/475,947A
;/ CURRENT FILING DATE: 1999-12-31
;/ NUMBER OF SEQ ID NOS: 346
;/ SOFTWARE: Patentln Ver. 2.1
;/ SEQ ID NO 246
;/ LENGTH: 60
;/ TYPE: DNA
;/ ORGANISM: human
US-09-475-947A-246

Query Match 40.0%; Score 24; DB 4; Length 60;
Best Local Similarity 68.8%; Pred. No. 13;
Matches 33; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 7 ACAGCCGCCACCCGACATGCGCTCAGAGCGCTCCGGATACGAGACCGGCC 54
Db 13 ACCGCCCCCAAGCCAGCGGTGTACCTCGGCCCGGACACGAGCGGCC 60

RESULT 10
US-09-489-039A-2790/c
;/ Sequence 2790, Application US/09489039A
;/ Patent No. 6610836
;/ GENERAL INFORMATION:
;/ APPLICANT: Gary Breton et. al
;/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
;/ TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
;/ FILE REFERENCE: 2709.2004001
;/ CURRENT APPLICATION NUMBER: US/09/489,039A
;/ CURRENT FILING DATE: 2000-01-27
;/ PRIOR APPLICATION NUMBER: US 60/117,747
;/ PRIOR FILING DATE: 1999-01-29
;/ NUMBER OF SEQ ID NOS: 14342
;/ SEQ ID NO 2790
;/ LENGTH: 1845
;/ TYPE: DNA
;/ ORGANISM: Klebsiella pneumoniae
US-09-489-039A-2790

Query Match 40.0%; Score 24; DB 4; Length 1845;
Best Local Similarity 64.3%; Pred. No. 18;
Matches 36; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

Qy 5 GTACAGCCGCCACCCGACATGCGCTCAGAGCGCTCCGGATACGAGACCGGCCCT 60
Db 1363 GGACAGCGTACCCCGCAAAATGACAGCGGTACTTAAGTACGATCCGACACT 1308

RESULT 11
US-09-489-039A-2874
;/ Sequence 2874, Application US/09489039A
;/ Patent No. 6610836
;/ GENERAL INFORMATION:
;/ APPLICANT: Gary Breton et. al
;/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
;/ TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
;/ FILE REFERENCE: 2709.2004001

/ CURRENT APPLICATION NUMBER: US/09/489,039A
/ CURRENT FILING DATE: 2000-01-27
/ PRIOR APPLICATION NUMBER: US 60/117,747
/ PRIOR FILING DATE: 1999-01-29
/ NUMBER OF SEQ ID NOS: 14342
/ SEQ ID NO 2874
/ LENGTH: 2397
/ TYPE: DNA
/ ORGANISM: Klebsiella pneumoniae
/ US-09-489-039A-2874

Query Match

Best Local Similarity 40.8%; Score 24; DB 4; Length 2397;
Matches 36; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

Qy 5 GTACAGCCGACCGGACATGCGGTGACGAGCGCTCCGATTCGAGACCGGCGCT 60
Db 1083 GACAGCGTACCGGACGAAATGACGAGGTGAATCTTAATGATACCGACACCT 1138

RESULT 12

US-09-105-537-40
/ Sequence 40, Application US/09105537A
/ Patent No. 6265202
/ GENERAL INFORMATION:
/ APPLICANT: Sherman, D.H.
/ APPLICANT: Liu, H.
/ APPLICANT: Xue, Y.
/ APPLICANT: Zhao, L.
/ TITLE OF INVENTION: DNA encoding methymycin and pikromycin
/ FILE REFERENCE: 600,438U1
/ CURRENT APPLICATION NUMBER: US/09/105,537A
/ CURRENT FILING DATE: 1998-06-26
/ NUMBER OF SEQ ID NOS: 43
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 40
/ LENGTH: 2787
/ TYPE: DNA
/ ORGANISM: Streptomyces venezuelae
/ US-09-105-537-40

Query Match

Best Local Similarity 39.3%; Score 23.6; DB 3; Length 2787;
Matches 32; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 15 ACCCGCATGCGCTGACGAGCGCTCCGATACGAGACCGGCGCT 60
Db 1292 ACCCGCATGACGACCGGCGCTTCGACTCTTCGACCGGCTCT 1337

RESULT 13

US-09-320-878-21
/ Sequence 21, Application US/09320878A
/ Patent No. 6117659
/ GENERAL INFORMATION:
/ APPLICANT: ASHLEY, Gary
/ APPLICANT: BETLACH, Melanie C.
/ APPLICANT: BETLACH, Mary C.
/ APPLICANT: MCDANIEL, Robert
/ APPLICANT: TANG, Li
/ TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
/ FILE REFERENCE: 30622002120
/ CURRENT APPLICATION NUMBER: US/09/320,878A
/ CURRENT FILING DATE: 1999-05-27
/ EARLIER APPLICATION NUMBER: CIP OF 09/141,908
/ EARLIER FILING DATE: 1998-08-28
/ EARLIER APPLICATION NUMBER: CIP OF 09/073,538
/ EARLIER FILING DATE: 1998-05-06
/ EARLIER APPLICATION NUMBER: CIP OF 08/846,247
/ EARLIER FILING DATE: 1997-04-30
/ EARLIER APPLICATION NUMBER: 60/119,139
/ EARLIER FILING DATE: 1999-02-08
/ EARLIER APPLICATION NUMBER: 60/100,880

/ EARLIER FILING DATE: 1998-09-22
/ EARLIER APPLICATION NUMBER: 60/087,080
/ EARLIER FILING DATE: 1998-05-28
/ NUMBER OF SEQ ID NOS: 34
/ SOFTWARE: Patentln Ver. 2.0
/ SEQ ID NO 21
/ LENGTH: 5970
/ TYPE: DNA
/ ORGANISM: Streptomyces venezuelae
/ US-09-320-878-21

Query Match

Best Local Similarity 39.3%; Score 23.6; DB 3; Length 5970;
Matches 32; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 15 ACCCGCATGCGCTGACGAGCGCTCCGATACGAGACCGGCGCT 60
Db 4030 ACCCGCATGACGACCGGCGCTTCGACTCTTCGACCGGCTCT 4075

RESULT 14

US-09-141-908-11
/ Sequence 11, Application US/09141908
/ Patent No. 6503741
/ GENERAL INFORMATION:
/ APPLICANT: ASHLEY, Gary
/ APPLICANT: BETLACH, Melanie C.
/ APPLICANT: BETLACH, Mary C.
/ APPLICANT: MCDANIEL, Robert
/ APPLICANT: TANG, Li
/ TITLE OF INVENTION: Combinatorial Polyketide Libraries Produced Using a
/ FILE REFERENCE: 30622002100
/ CURRENT APPLICATION NUMBER: US/09/141,908
/ CURRENT FILING DATE: 1998-08-28
/ EARLIER APPLICATION NUMBER: CIP OF 09/073,538
/ EARLIER FILING DATE: 1998-05-06
/ EARLIER APPLICATION NUMBER: CIP OF 08/846,247
/ EARLIER FILING DATE: 1997-04-30
/ EARLIER APPLICATION NUMBER: PROV. 60/076,919
/ EARLIER FILING DATE: 1998-03-05
/ EARLIER APPLICATION NUMBER: PROV. 60/087,080
/ EARLIER FILING DATE: 1998-05-28
/ NUMBER OF SEQ ID NOS: 31
/ SOFTWARE: Patentln Ver. 2.0
/ SEQ ID NO 11
/ LENGTH: 5970
/ TYPE: DNA
/ ORGANISM: Streptomyces venezuelae
/ US-09-141-908-11

Query Match

Best Local Similarity 39.3%; Score 23.6; DB 4; Length 5970;
Matches 32; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 15 ACCCGCATGCGCTGACGAGCGCTCCGATACGAGACCGGCGCT 60
Db 4030 ACCCGCATGACGACCGGCGCTTCGACTCTTCGACCGGCTCT 4075

RESULT 15

US-09-657-440-21
/ Sequence 21, Application US/09657440
/ Patent No. 6509455
/ GENERAL INFORMATION:
/ APPLICANT: ASHLEY, Gary
/ APPLICANT: BETLACH, Melanie C.
/ APPLICANT: BETLACH, Mary C.
/ APPLICANT: MCDANIEL, Robert
/ APPLICANT: TANG, Li
/ TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
/ FILE REFERENCE: 30622002120
/ CURRENT APPLICATION NUMBER: US/09/657,440

; CURRENT FILING DATE: 2000-09-07
; PRIOR APPLICATION NUMBER: 09/320,878
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: CIP OF 09/141,908
; PRIOR FILING DATE: 1998-08-28
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
; LENGTH: 5970
; TYPE: DNA
; ORGANISM: Streptomyces venezuelae
US-09-657-440-21

Query Match 39.1%; Score 23.6; DB 4; Length 5970;
Best Local Similarity 69.6%; Pred. No. 27;
Matches 32; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 15 ACCCGCAGATGGCGCTCAGAGCGGCTCCGATACGAGACCGGCGCT 60
DB 4030 ACCCGCAGATGAGAGCGGCGGCTCGCACTTTGAGCGGCTCCT 4075

Search completed: January 16, 2005, 03:02:00
Job time : 44.2 secs

This Page Blank (uspto)

Oy 7 ACAGCGCCACCCGACATGCGCTCAGCAGCGCTCCGAGATACGAGCCGCGCC 59
 Db 250 ACCGCCCCCAGCCACGCTGTCACCTCGGCTCCGAGACGAGCCGCGCC 198
 RESULT 2
 BUI152566
 LOCUS
 DEFINITION BUI152566 1349 bp mRNA linear EST 03-SEP-2002
 5', mRNA sequence.
 ACCESSION BUI152566
 VERSION BUI152566.1 GI:22666098
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 1349)
 NIH-MGC <http://mgs.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: DCTD/DRP
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: L1CM2569 row: 6 column: 16
 High quality sequence stop: 291.
 Location/Qualifiers
 1..1349
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6380559"
 /tissue_type="carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_id="NIH_MGC_40"
 /note="Organ: prostate; Vector: pORF7; Site:1; XhoI;
 Site 2: EcoRI; cDNA made by oligo-dT priming.
 directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCACGAG(G). Library constructed by
 Ling Hong in the laboratory of Gerald M. Rubin (University
 of California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC library."

ORIGIN
 Query Match 48.3%; Score 29; DB 5; Length 1349;
 Best Local Similarity 71.7%; Pred. No. 18;
 Matches 38; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
 Oy 7 ACAGCGCCACCCGACATGCGCTCAGCAGCGCTCCGAGATACGAGCCGCGCC 59
 Db 265 ACCGCCCCCAGCCACGCTGTCACCTCGGCTCCGAGACGAGCCGCGCC 317

RESULT 3
 BUI256331 486 bp mRNA linear EST 10-DEC-2002
 LOCUS BUI256331
 DEFINITION BUI256331 RIKEN full-length enriched, visual cortex Mus musculus
 cDNA clone K330023N10 5', mRNA sequence.
 ACCESSION BUI256331
 VERSION BUI256331.1 GI:26437843
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 486)

REFERENCE
 1 (bases 1 to 486)

AUTHORS
 Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
 Nikaado, I., Osato, N., Saito, R., Suzuki, H., Yamaguchi, Y., Nogami, A.,
 Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nishimura, T.,
 Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bul, C.,
 Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H.,
 Batilov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V.,
 Chochia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A.,
 Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T.,
 Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,
 Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A.,
 Kawaji, H., Kawasawa, Y., Kedzierzki, R.M., King, B.L., Kongaya, A.,
 Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R.,
 Maltas, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,
 Numata, K., Okita, T., Pavan, W.J., Petrea, G., Pesole, G.,
 Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S.,
 Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M.,
 Sadelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K.,
 Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M.,
 Verardo, R., Wagner, L., Wahlestedt, C., Wang, X., Watanabe, Y.,
 Wells, C., Wilming, L.G., Wyrshaw-Boris, A., Yanagisawa, M., Yang, I.,
 Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,
 Hayatsu, N., Hirozane-Kishikawa, T., Kono, H., Nakamura, M.,
 Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
 Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y.,
 Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,
 Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S.,
 Rogers, J., Birney, E. and Hayashizaki, Y.
 Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 JOURNAL
 MEDLINE
 PUBMED
 22354683
 12466851
 COMMENT
 Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suenhiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gscc.riken.jp, URL:http://genome.gsc.riken.jp/
 Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S.,
 Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Kono, H.,
 Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R.,
 Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K.,
 Shiraki, T., Tagami, M., Waki, K., Watanabe, Y., Watanabe, M., and
 Hayashizaki, Y. Direct Submission
 Computational Analysis of Full-Length Mouse cDNAs Compared with
 Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multichannel sequencer. Genome Res.
 10 (11), 1757-1771 (2000)
 Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.
 Tissues were provided by Michela Fagiolini and Takao K. Hensch (
 Laboratory for Neuronal Circuit Development Brain Science Institute
 RIKEN 2-1 Hirosawa, Mako-shi, Saitama 351-0198 Japan) whose
 assistance we gratefully acknowledge. Please visit our web site
 (<http://genome.gsc.riken.go.jp>) for further details.
 Location/Qualifiers
 1..486
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"

ORIGIN
/clone="K330023N10"
/tissue_type="visual cortex"
/clone_lib="RIKEN full-length enriched, visual cortex"

Query Match 46.0%; Score 27.6; DB 5; Length 486;
Best Local Similarity 66.1%; Pred. No. 57;
Matches 39; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 2 GATGATACAGCCGACATGCGCTCAGCGCCCTCGGATACGAGACCGCGCCCT 60
123 GATGCGCGCGCGCGCCCTTACGTCCTACGCGCGGAGAAAGAACCGCGCCCT 65

RESULT 4
BM791359 604 bp mRNA linear EST 05-MAR-2002
LOCUS K-EST0071342 S21SNUS20 Homo sapiens CDNA clone S21SNUS20-14-A06 5',
DEFINITION mRNA sequence.

ACCESSION BM791359
VERSION BM791359.1 GI:19139591
KEYWORDS EST.

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 604)
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.

TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS

Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 14 row: A column: 06
High quality sequence stop: 604.
Location/Qualifiers
1. 604

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S21SNUS20-14-A06"
/sex="F"
/tissue_type="Stomach"
/cell_type="Floating aggregates"
/cell_line="SNU-520"
/lab_host="Top10F"
/clone_lib="S21SNUS20"
/note="Organ: Stomach; Vector: pTZ189P; Site 1: EcoRI;
Site 2: NotI; The poly (A)+ RNA was: dephosphorylated with
bacterial alkaline phosphatase (BAP) and then deacapped
with tabacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including EcoR
I site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized from oligo dt-selected mRNA by
priming with dt-tailed vector. The dt-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F' by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."

ORIGIN

Query Match 45.7%; Score 27.4; DB 4; Length 604;
Best Local Similarity 69.8%; Pred. No. 65;

Matches 37; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
QY 7 ACAGCGCACCCGACATGCGCTCAGAGCGCTCCGATACGAGACCGCGGCC 59
34 ACCGACCCCGACGCCGATGTACCTCGCGCCCGAGACACGCGCGGCC 86

RESULT 5
BUS42454 877 bp mRNA linear EST 13-SEP-2002
LOCUS AGENCOURT_10322173 NIH_MGC_40 Homo sapiens CDNA clone IMAGE:5574322
DEFINITION 5', mRNA sequence.

ACCESSION BUS42454
VERSION BUS42454.1 GI:22852937
KEYWORDS EST.

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 877)
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
NIH-MGC <http://mgi.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: DCTD/DRP

CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LINL at:
<http://image.jnl.gov>
Plate: LUCM2769 row: h column: 02
High quality sequence stop: 760.
Location/Qualifiers
1. 877

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5574322"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
/note="Organ: prostate; Vector: pORF7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Iing Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC library."

ORIGIN

Query Match 45.7%; Score 27.4; DB 5; Length 877;
Best Local Similarity 69.8%; Pred. No. 64;
Matches 37; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 7 ACAGCGCACCCGACATGCGCTCAGAGCGCTCCGATACGAGACCGCGGCC 59
41 ACCGCGCCCGACGCCGATGTACCTCGCGCCCGAGACACGCGCGGCC 93

RESULT 6
CA489836 959 bp mRNA linear EST 14-NOV-2002
LOCUS AGENCOURT_10810668 MAPcl Homo sapiens CDNA clone IMAGE:6722324 5',
DEFINITION mRNA sequence.

ACCESSION CA489836
VERSION CA489836.1 GI:24952627
KEYWORDS EST.

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 959)
 AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Kristi A. Eglund, Ira Pastan
 CDNA Library Preparation: Invitrogen Corp
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
 Plate: L1AM14284 row: n column: 20
 High quality sequence start: 31
 High quality sequence stop: 446.
 Location/Qualifiers
 1..959

FEATURES
 source
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:672324"
 /cell_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231, HEBT-HME1, LNCaP"
 /lab_host="EMDH10B"
 /clone_1ib="MAPCL"
 /note="Vector: PCMV-SPORT6; Site_1: EcoRV; Site_2: Not I; Subtracted with brain, liver, lung, kidney and muscle. Directionally cloned. Priming method: oligo-dT. Average insert size: 1800 bp. Library amplification: 26,000 fold. Kristi A. Eglund, James J. Vincent, Robert Strausberg, Bungkok Lee & Ira Pastan: Discovery of new breast cancer genes encoding membrane and secreted proteins. Manuscript submitted."

ORIGIN

Query Match 45.7%; Score 27.4; DB 6; Length 959;
 Best Local Similarity 69.8%; Pred. No. 64;
 Matches 37; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 7 ACAGGCCACCCGACATGCGTCACGAGCGCTCCGGATCGAGACCGGGCC 59
 DB 355 ACCGCCCCCGACGCCAGGTGTCCTCGGCCCCGACACGAGCGCGCC 407

RESULT 7
 LOCUS BU148487 1113 bp mRNA linear EST 03-SEP-2002
 DEFINITION AGENCOURT 8670479 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6380642
 ACCESSION BU148487
 VERSION BU148487.1 GI:22662019
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 1113)
 NIH-MGC <http://mgs.nci.nih.gov/>.
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
 COMMENT Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: DCTD/DRP
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
 Plate: L1CM2559 row: j column: 03
 High quality sequence stop: 235.

FEATURES

source

ORIGIN

Query Match 45.7%; Score 27.4; DB 5; Length 1113;
 Best Local Similarity 69.8%; Pred. No. 63;
 Matches 37; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 7 ACAGGCCACCCGACATGCGTCACGAGCGCTCCGGATCGAGACCGGGCC 59
 DB 36 ACCGCCCCCGACGCCAGGTGTCCTCGGCCCCGACACGAGCGCGCC 88

RESULT 8

LOCUS BU542996

DEFINITION

AGENCOURT 10338707 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6574979
 5', mRNA sequence.

ACCESSION

BU542996

VERSION

BU542996.1

KEYWORDS

EST.

SOURCE

Homo sapiens

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 1130)
 NIH-MGC <http://mgs.nci.nih.gov/>.
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
 COMMENT Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: DCTD/DRP
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
 Plate: L1CM2771 row: c column: 11
 High quality sequence start: 27
 High quality sequence stop: 246.
 Location/Qualifiers
 1..1130

FEATURES

source

/organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6574979"
 /tissue_type="carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_1ib="NIH_MGC_40"
 /note="Organ: prostate; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 45.7%; Score 27.4; DB 5; Length 1130;
Best Local Similarity 69.8%; Pred. No. 63;
Matches 37; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 7 ACAGCGCCACCCGACATGCGCTCAGAGCGCTCCGATACGAGACCGCGCC 59
DB 75 ACCGCCCCCGCCAGCCGTCCTCCTCGCGCCCGGACACGCGCGCC 127

RESULT 9
BO936898 1234 bp mRNA linear EST 21-ANG-2002
LOCUS AGENCOURT_8919110 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6383855
DEFINITION 5', mRNA sequence.

ACCESSION BO936898
VERSION BO936898
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE NIH-MGC http://img.mgc.nci.nih.gov/
JOURNAL 1 (bases 1 to 1234)
COMMENT Unpublished (1999) National Institutes of Health, Mammalian Gene Collection (MGC)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
plate: LLCM2577 row: o column: 24
High quality sequence stop: 245.

FEATURES

source

Location/Qualifiers

1..1234
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6383855"
/feature_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_1lb="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN

Query Match 45.7%; Score 27.4; DB 5; Length 1234;
Best Local Similarity 69.8%; Pred. No. 63;
Matches 37; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 7 ACAGCGCCACCCGACATGCGCTCAGAGCGCTCCGATACGAGACCGCGCC 59
DB 36 ACCGCCCCCGCCAGCCGTCCTCCTCGCGCCCGGACACGCGCGCC 88

RESULT 10
BO943554 1268 bp mRNA linear EST 21-ANG-2002
LOCUS AGENCOURT_8777815 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6384308
DEFINITION 5', mRNA sequence.
ACCESSION BO943554
VERSION BO943554.1 GI:22359032

KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE NIH-MGC http://img.mgc.nci.nih.gov/
JOURNAL 1 (bases 1 to 1268)
COMMENT Unpublished (1999) National Institutes of Health, Mammalian Gene Collection (MGC)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
plate: LLCM2579 row: b column: 21
High quality sequence stop: 177.

FEATURES

source

Location/Qualifiers

1..1268
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6384308"
/feature_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_1lb="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN

Query Match 45.7%; Score 27.4; DB 5; Length 1268;
Best Local Similarity 69.8%; Pred. No. 63;
Matches 37; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 7 ACAGCGCCACCCGACATGCGCTCAGAGCGCTCCGATACGAGACCGCGCC 59
DB 96 ACCGCCCCCGCCAGCCGTCCTCCTCGCGCCCGGACACGCGCGCC 148

RESULT 11
BO920055 1343 bp mRNA linear EST 20-ANG-2002
LOCUS AGENCOURT_10031674 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6481609
DEFINITION 5', mRNA sequence.

ACCESSION BO920055
VERSION BO920055
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE NIH-MGC http://img.mgc.nci.nih.gov/
JOURNAL 1 (bases 1 to 1343)
COMMENT Unpublished (1999) National Institutes of Health, Mammalian Gene Collection (MGC)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov

Plate: L16M2663 row: e column: 02
 High quality sequence start: 56
 High quality sequence stop: 237.
 Location/Qualifiers

FEATURES

source

1.1343

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:6481609"

/issue_type="carcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC 40"

/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming; Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 45.7%; Score 27.4; DB 5; Length 1343;
 Best Local Similarity 69.8%; Pred. No. 63;
 Matches 37; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 7 ACAGCGCCACCGGCATGCGTCAAGAGCGCTCCGATATACAGACCGCGGCC 59
 DB 78 ACCGCCCCCAGCCACCGTGTCACTCTGCCCCGACACACGCGGCC 130

RESULT 12

LOCUS

DEFINITION

AGENCOURT_10334841 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6574725

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1.1420

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:6574725"

/issue_type="carcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC 40"

/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library."

ORIGIN

Query Match 45.7%; Score 27.4; DB 5; Length 1420;
 Best Local Similarity 69.8%; Pred. No. 63;
 Matches 37; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 7 ACAGCGCCACCGGCATGCGTCAAGAGCGCTCCGATATACAGACCGCGGCC 59
 DB 25 ACCGCCCCCAGCCACCGTGTCACTCTGCCCCGACACACGCGGCC 77

RESULT 13

LOCUS

DEFINITION

AGENCOURT_10327072 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6575349

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1.1531
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6575349"
 /issue_type="carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC 40"
 /note="Organ: prostate; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 45.7%; Score 27.4; DB 5; Length 1531;
 Best Local Similarity 69.8%; Pred. No. 62;
 Matches 37; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 7 ACAGCGCCACCGGCATGCGTCAAGAGCGCTCCGATATACAGACCGCGGCC 59
 DB 36 ACCGCCCCCAGCCACCGTGTCACTCTGCCCCGACACACGCGGCC 88

RESULT 14

LOCUS

DEFINITION

W16262 451 bp mRNA linear EST 12-SEP-1996
 mb64c07.t1 Soares mouse p3NMFI9.5 Mus musculus cDNA clone
 IMAGE:334188 5' similar to gb:228407 60S RIBOSOMAL PROTEIN L8

This Page Blank (uspto)

XX Sequence 60 BP; 12 A; 22 C; 19 G; 7 T; 0 U; 0 Other;
SQ

Query Match 100.0%; Score 60; DB 2; Length 60;
Best Local Similarity 100.0%; Pred. No. 3,6e-11;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTAAGTACAGCGCCACCCGACATGGCGTCCGAGCGCTCCGATACGAGACCGGCGCCT 60
Db 1 GGTAAGTACAGCGCCACCCGACATGGCGTCCGAGCGCTCCGATACGAGACCGGCGCCT 60

RESULT 2
AAV48322

ID AAV48322 standard; DNA; 60 BP.

XX AAV48322;

DT 20-NOV-1998 (first entry)

DE Nucleotide sequence encoding MUC1 tandem repeat unit R7.

KM ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
XX tumour; tumour-associated antigen.

OS Homo sapiens.

PN WO9837095-A2.

PD 27-AUG-1998.

PF 24-FEB-1998; 98WO-US003693.

PR 24-FEB-1997; 97US-0038253P.

XX (THER-) THERION BIOLOGICS CORP.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

PI (DAND) DANA FARBER CANCER INST INC.

XX Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;

DR WPI; 1998-467492/40.

PT New recombinant pox virus for tumour therapy - comprises DNA encoding an
immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.

PS Disclosure; Page 11; 42pp; English.

CC The MUC1 tandem repeat units AAV48317-V48325 were used to create an
immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
CC (RPV). The RPV was used in a pharmaceutical composition also containing
an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The
recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
CC does not undergo significant genetic deletion, thereby providing an
unexpectedly stable and immunogenic pox virus. They can be used to
prevent or treat tumours expressing MUC1 tumour-associated antigens

XX Sequence 60 BP; 14 A; 24 C; 16 G; 6 T; 0 U; 0 Other;

Query Match 63.7%; Score 38.2; DB 2; Length 60;
Best Local Similarity 78.0%; Pred. No. 0.0011;
Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 1 GGTAAGTACAGCGCCACCCGACATGGCGTCCGAGCGCTCCGATACGAGACCGGCGCCT 59
Db 1 GGTAAGTACAGCGCTCCACCTGCACACGCGGTCCACAGCGCGCCAGACATCTGCGCGCC 59

RESULT 3
AAV48317

ID AAV48317 standard; DNA; 60 BP.

XX AAV48317;

XX 20-NOV-1998 (first entry)

DE Nucleotide sequence encoding MUC1 tandem repeat unit R2.

KM ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
XX tumour; tumour-associated antigen.

OS Homo sapiens.

PN WO9837095-A2.

PD 27-AUG-1998.

PF 24-FEB-1998; 98WO-US003693.

PR 24-FEB-1997; 97US-0038253P.

XX (THER-) THERION BIOLOGICS CORP.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

PI (DAND) DANA FARBER CANCER INST INC.

XX Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;

DR WPI; 1998-467492/40.

PT New recombinant pox virus for tumour therapy - comprises DNA encoding an
immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.

PS Disclosure; Page 11; 42pp; English.

CC The MUC1 tandem repeat units AAV48317-V48325 were used to create an
immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
CC (RPV). The RPV was used in a pharmaceutical composition also containing
an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The
recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
CC does not undergo significant genetic deletion, thereby providing an
unexpectedly stable and immunogenic pox virus. They can be used to
prevent or treat tumours expressing MUC1 tumour-associated antigens

XX Sequence 60 BP; 17 A; 21 C; 13 G; 9 T; 0 U; 0 Other;

Query Match 62.7%; Score 37.6; DB 2; Length 60;
Best Local Similarity 76.7%; Pred. No. 0.0018;
Matches 46; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 1 GGTAAGTACAGCGCCACCCGACATGGCGTCCGAGCGCTCCGATACGAGACCGGCGCCT 60
Db 1 GGCAAGTACAGCGCCACCCGACATGGCGTCCGATACGAGACCTGATCAAGACCTGCACTT 60

RESULT 4
AAV48321

ID AAV48321 standard; DNA; 60 BP.

XX AAV48321;

DT 20-NOV-1998 (first entry)

DE Nucleotide sequence encoding MUC1 tandem repeat unit R6.

KM ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
XX tumour; tumour-associated antigen.

OS Homo sapiens.

PN WO9837095-A2.

PD 27-AUG-1998.

PF 24-FEB-1998; 98WO-US003693.

PR 24-FEB-1997; 97US-0038253P.

```

XX (THER-) THERION BIOLOGICS CORP.
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA (DAND) DANA FARBEN CANCER INST INC.
XX
XX Schlom J, Kantor J, Kufe D, Panicali D, Gritz L,
DR WPI; 1998-467492/40.
XX
XX New recombinant pox virus for tumour therapy - comprises DNA encoding an
PT immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
XX
XX Disclosure; Page 11; 42pp; English.
XX
XX The MUC1 tandem repeat units AAV48317-V48325 were used to create an
CC immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
CC (RPV). The RPV was used in a pharmaceutical composition also containing
CC an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The
CC recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
CC does not undergo significant genetic deletion, thereby providing an
CC unexpectedly stable and immunogenic pox virus. They can be used to
CC prevent or treat tumours expressing MUC1 tumour-associated antigens
XX
SQ Sequence 60 BP; 13 A; 27 C; 17 G; 3 T; 0 U; 0 Other;

Query Match          61.0%; Score 36.6; DB 2; Length 60;
Best Local Similarity 76.3%; Pred. No. 0.0039;
Matches 45; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

OY 1 GGTAGTACAGCGCCACCGCATGCGCTCAGAGCGCTCCGATACGAGACCGCGCC 59
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1 GGCACACCGCACCGCCCGGCACACGGGGTCAACAAGCGCCCGACAGACTGACTGCGCC 59

RESULT 5
AAD00385
ID AAD00385 strand: DNA; 525 BP.
XX
XX AAD00385;
XX
XX 29-AUG-2000 (first entry)
XX
XX Human Mucin 1 (MUC-1) protein fragment encoding DNA #2.
XX
XX Human; Mucin 1; MUC-1; tumour; pMR30 expression vector; anti-tumour;
XX therapy; immune response; cytostatic; vaccine; ds.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 1..525
XX FT /*tag= a
XX FT /product= "MUC-1 protein fragment"
XX
XX WO200025827-A2.
XX
XX 11-MAY-2000.
XX
XX 18-OCT-1999; 99WO-EP007874.
XX
XX 30-OCT-1998; 98IT-MI002330.
XX
XX (MENA ) MENARINI RICERCHE SPA.
XX
XX Parente D, Di Massimo AM, De Santis R;
XX
XX WPI; 2000-365410/31.
XX
XX P-PSDB; AAV71021.
XX
XX Composition containing one or more DNA molecules encoding fragments of a
PT Mucin 1 (MUC-1) protein overexpressed in tumor cells, useful in anti-
XX tumor therapy.

```

```

PS Claim 16; Fig 2; 56pp; English.
XX
XX The present sequence is a DNA encoding a fragment of human Mucin 1 (MUC-
CC 1) antigenic protein which is overexpressed in tumour cells. The sequence
CC was obtained from BR20 tumour cells by reverse transcriptase-PCR and
CC corresponds to nucleotides 205-720 of the EMBL sequence J05581 with a
CC start codon and two stop codons. The present sequence is cloned into a
CC pMR30 expression vector and used in pharmaceutical composition e.g.
CC vaccine for inducing an antigen-specific anti-tumour immune response.
CC Composition containing this DNA molecule is useful in anti-tumour therapy
CC of patients affected with tumours characterised by high MUC-1 expression
XX
SQ Sequence 525 BP; 111 A; 199 C; 128 G; 87 T; 0 U; 0 Other;

Query Match          61.0%; Score 36.6; DB 3; Length 525;
Best Local Similarity 76.3%; Pred. No. 0.0046;
Matches 45; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

OY 1 GGTAGTACAGCGCCACCGCATGCGCTCAGAGCGCTCCGATACGAGACCGCGCC 59
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 256 GGAAGTACTGCTCCACACGACACCGGTGTTACTCGGCTCCGATACCGAGCCGCCCC 314

RESULT 6
AAD00391
ID AAD00391 strand: DNA; 891 BP.
XX
XX AAD00391;
XX
XX 15-SEP-2003 (revised)
XX
XX 29-AUG-2000 (first entry)
XX
XX Ubiquitin-E. coli LacI-human Mucin 1 fusion protein encoding DNA #2.
XX
XX Ubiquitin; LacI; beta-galactosidase; fusion protein; human; Mucin 1;
XX MUC-1; tumour; pMR30 expression vector; anti-tumour; therapy;
XX immune response; cytostatic; vaccine; ds.
XX
XX Homo sapiens.
XX
XX Escherichia coli.
XX
XX Chimeric.
XX
XX Key Location/Qualifiers
XX CDS 1..891
XX FT /*tag= a
XX FT /product= "Ubiquitin-LacI-MUC-1 fusion protein"
XX
XX misc_feature 1..369
XX FT /*tag= b
XX FT /label= UBILacI DNA
XX FT /note= "Includes ubiquitin-E. coli LacI fusion DNA"
XX
XX misc_feature 370..891
XX FT /*tag= c
XX FT /note= "Human MUC-1 partial DNA that corresponds to
XX nucleotides 205-720 of the EMBL sequence J05581 with two
XX stop codons"
XX
XX WO200025827-A2.
XX
XX 11-MAY-2000.
XX
XX 18-OCT-1999; 99WO-EP007874.
XX
XX 30-OCT-1998; 98IT-MI002330.
XX
XX (MENA ) MENARINI RICERCHE SPA.
XX
XX Parente D, Di Massimo AM, De Santis R;
XX
XX WPI; 2000-365410/31.
XX
XX P-PSDB; AAV71027.
XX
XX Composition containing one or more DNA molecules encoding fragments of a
PT Mucin 1 (MUC-1) protein overexpressed in tumor cells, useful in anti-

```

```

PT tumor therapy.
XX
PS Claim 18; Fig 8; 56bp; English.
XX
CC The present sequence is a DNA encoding a fusion protein consisting of
CC human Mucin 1 (MUC-1) fragment fused to UBIlacti sequence at the N-
CC terminus. The UBIlacti sequence consists of ubiqlutin from MCF7 cell line
CC and a portion of E. coli beta-galactosidase (laci). MUC-1 is an antigenic
CC protein overexpressed in tumour cells. The present sequence is cloned
CC into a pMS30 expression vector and used in pharmaceutical composition
CC e.g. vaccine for inducing an antigen-specific anti-tumour immune
CC response. Composition containing this DNA molecule is useful in anti-
CC tumour therapy of patients affected with tumours characterised by high
CC MUC-1 expression. (Updated on 15-SEP-2003 to standardise OS field)
XX
SQ Sequence 891 BP; 203 A; 307 C; 225 G; 156 T; 0 U; 0 Other;

Query Match
Best Local Similarity 61.0%; Score 36.6; DB 3; Length 891;
Matches 45; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 GGTAGTACAGCGCCACCGCCGATGCGCTCAGAGCGCTCGGATACGAGCCGCGCC 59
Db 622 GGAAGTACTCTCTCCACCAAGACGCTGTACTCTCGGCTCGGATACGAGCCGCGCC 680

RESULT 7
AAD00388
ID AAD00388 standard; DNA; 1371 BP.
AC AAD00388;
XX
DT 29-AUG-2000 (first entry)
XX
DE Human Mucin 1 (MUC-1) protein fragment encoding DNA #5.
XX
KW Human; Mucin 1; MUC-1; tumour; pMS30 expression vector; anti-tumour;
XX therapy; immune response; cytostatic; vaccine; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT 1..1371
FT /*tag= a
FT /*product= "MUC-1 protein fragment"
FT
FT
FT
XX
PN WO200025827-A2.
XX
PD 11-MAY-2000.
XX
PF 18-OCT-1999; 99WO-EP007874.
XX
PR 30-OCT-1998; 98IT-MI002330.
XX
PA (MENA ) MENARINI RICERCH SPA.
XX
PI Parente D, Di Massimo AM, De Santis R;
XX
DR WPI; 2000-365410/31.
DR P-PSDB; AAY71024.
XX
XX
PT Composition containing one or more DNA molecules encoding fragments of a
PT Mucin 1 (MUC-1) protein overexpressed in tumor cells, useful in anti-
PT tumor therapy.
XX
PS Claim 16; Fig 5; 56bp; English.
XX
CC The present sequence is a DNA encoding a fragment of human Mucin 1 (MUC-
CC 1) antigenic protein which is overexpressed in tumour cells. The sequence
CC was obtained by PCR from plasmids pMS167, pMS167, pMS168 and pMS169
CC which contain MUC-1 DNA from BT20 tumour cells. It corresponds to
CC nucleotides 136-1497 of the EMBL sequence J05581 with a start codon and
CC two stop codons. The present sequence is cloned into a pMS30 expression

```

```

CC vector and used in pharmaceutical composition e.g. vaccine for inducing
CC an antigen-specific anti-tumour immune response. Composition containing
CC this DNA molecule is useful in anti-tumour therapy of patients affected
CC with tumours characterised by high MUC-1 expression
XX
SQ Sequence 1371 BP; 300 A; 450 C; 321 G; 300 T; 0 U; 0 Other;

Query Match
Best Local Similarity 76.3%; Score 36.6; DB 3; Length 1371;
Matches 45; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 GGTAGTACAGCGCCACCGCCGATGCGCTCAGAGCGCTCGGATACGAGCCGCGCC 59
Db 325 GGAAGTACCGCTCCACCAAGACGCTGTACTCTCGGCTCGGATACGAGCCGCGCC 383

RESULT 8
AAD00394
ID AAD00394 standard; DNA; 1737 BP.
AC AAD00394;
XX
DT 15-SEP-2003 (revised)
DT 29-AUG-2000 (first entry)
XX
DE Ubiqlutin-E. coli Laci-human Mucin 1 fusion protein encoding DNA #5.
XX
KW Ubiqlutin; laci; beta-galactosidase; fusion protein; human; Mucin 1;
XX MUC-1; tumour; pMS30 expression vector; anti-tumour; therapy;
XX immune response; cytostatic; vaccine; ds.
XX
OS Homo sapiens.
XX
OS Escherichia coli.
XX
FH Key Location/Qualifiers
FT 1..1737
FT /*tag= a
FT /*product= "Ubiqlutin-Laci-MUC-1 fusion protein"
FT 1..369
FT /*tag= b
FT /*label= UBIlacti_DNA
FT /*note= "includes ubiqlutin-E. coli Laci fusion DNA"
FT 370..1737
FT /*tag= c
FT /*note= "Human MUC-1 partial DNA that corresponds to
FT nucleotides 136-1497 of the EMBL sequence J05581 with two
FT stop codons"
FT
FT
FT
XX
PN WO200025827-A2.
XX
PD 11-MAY-2000.
XX
PF 18-OCT-1999; 99WO-EP007874.
XX
PR 30-OCT-1998; 98IT-MI002330.
XX
PA (MENA ) MENARINI RICERCH SPA.
XX
PI Parente D, Di Massimo AM, De Santis R;
XX
DR WPI; 2000-365410/31.
DR P-PSDB; AAY71030.
XX
XX
PT Composition containing one or more DNA molecules encoding fragments of a
PT Mucin 1 (MUC-1) protein overexpressed in tumor cells, useful in anti-
PT tumor therapy.
XX
PS Claim 18; Fig 11; 56bp; English.
XX
CC The present sequence is a DNA encoding a fusion protein consisting of
CC human Mucin 1 (MUC-1) fragment fused to UBIlacti sequence at the N-
CC terminus. The UBIlacti sequence consists of ubiqlutin from MCF7 cell line

```

CC and a portion of E. coli beta-galactosidase (lacZ). MUC-1 is an antigenic
 CC protein overexpressed in tumour cells. The present sequence is cloned
 CC into a pMR810 expressed in tumour cells and used in pharmaceutical composition
 CC e.g. vaccine for inducing an antigen-specific anti-tumour immune
 CC response. Composition containing this DNA molecule is useful in anti-
 CC tumour therapy of patients affected with tumours characterised by high
 CC MUC-1 expression. (Updated on 15-SEP-2003 to standardise OS field)

XX
 SQ Sequence 1737 BP; 392 A; 558 C; 418 G; 369 T; 0 U; 0 Other;

Query Match 61.0%; Score 36.6; DB 3; Length 1737;
 Best Local Similarity 76.3%; Pred. No. 0.005;
 Matches 45; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 GGTAAGTACACCGCCGACATGCGCTGTCACGAGCGCTCCGATACGAGACCGGCGCC 59
 DB 691 GGAAATACCGCTCCACACGACGAGTGTACTCTCGGCTCCGATACGAGCGGCGCC 749

RESULT 9

AAV48320
 ID AAV48320 standard; DNA; 60 BP.

AC AAV48320;

DT 20-NOV-1998 (first entry)

DE Nucleotide sequence encoding MUC1 tandem repeat unit R5.

XX ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
 XX tumour; tumour-associated antigen.

OS Homo sapiens.

XX WO9837095-A2.

PD 27-AUG-1998.

PF 24-FEB-1998; 98WO-US003693.

PR 24-FEB-1997; 97US-0038253P.

PA (THER-) THERION BIOLOGICS CORP.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

PA (DAND) DANA FARBER CANCER INST INC.

PI Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;

XX WPI; 1998-467492/40.

PT New recombinant pox virus for tumour therapy - comprises DNA encoding an

XX immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.

PS Disclosure; Page 11; 42pp; English.

CC The MUC1 tandem repeat units AAV48317-V48325 were used to create an
 CC immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
 CC (RPV). The RPV was used in a pharmaceutical composition also containing
 CC an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The
 CC recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
 CC does not undergo significant genetic deletion, thereby providing an
 CC unexpectedly stable and immunogenic pox virus. They can be used to
 CC prevent or treat tumours expressing MUC1 tumour-associated antigens

XX
 SQ Sequence 60 BP; 9 A; 27 C; 15 G; 9 T; 0 U; 0 Other;

Query Match 54.7%; Score 32.8; DB 2; Length 60;

Best Local Similarity 71.7%; Pred. No. 0.079;

Matches 43; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 1 GGTAAGTACACCGCCGACATGCGCTGTCACGAGCGCTCCGATACGAGACCGGCGCT 60
 DB 1 GGTTCAACGCGCCCTGCTGTCACGATACATCCGCCGATACGAGCGGCGCT 60

RESULT 10

AAV48319
 ID AAV48319 standard; DNA; 60 BP.

XX AAV48319;

DT 20-NOV-1998 (first entry)

DE Nucleotide sequence encoding MUC1 tandem repeat unit R4.

XX ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
 XX tumour; tumour-associated antigen.

OS Homo sapiens.

XX WO9837095-A2.

PD 27-AUG-1998.

PF 24-FEB-1998; 98WO-US003693.

PR 24-FEB-1997; 97US-0038253P.

PA (THER-) THERION BIOLOGICS CORP.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

PA (DAND) DANA FARBER CANCER INST INC.

PI Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;

XX WPI; 1998-467492/40.

PT New recombinant pox virus for tumour therapy - comprises DNA encoding an

XX immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.

PS Disclosure; Page 11; 42pp; English.

CC The MUC1 tandem repeat units AAV48317-V48325 were used to create an
 CC immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
 CC (RPV). The RPV was used in a pharmaceutical composition also containing
 CC an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The
 CC recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
 CC does not undergo significant genetic deletion, thereby providing an
 CC unexpectedly stable and immunogenic pox virus. They can be used to
 CC prevent or treat tumours expressing MUC1 tumour-associated antigens

XX
 SQ Sequence 60 BP; 10 A; 23 C; 13 G; 14 T; 0 U; 0 Other;

Query Match 48.3%; Score 29; DB 2; Length 60;

Best Local Similarity 71.7%; Pred. No. 1.6;

Matches 38; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 7 ACAGCGCCACCGCCGACATGCGCTGTCACGAGCGCTCCGATACGAGACCGGCGCC 59
 DB 7 AAGGCTCTCCCGCTCATGCGGTTACTTGTCTCCAGATACGCGCCAGCTCC 59

RESULT 11

ABK36806
 ID ABK36806 standard; DNA; 72 BP.

XX ABK36806;

DT 08-MAY-2002 (first entry)

DE Human DNA encoding MUC1F segment 8.

XX Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus;

XX viral infection; human immunodeficiency virus; melanoma;

XX bacterial infection; Salmonella; paratuberculosis; parasitic infection;
 XX Trypanosoma; Toxoplasma; Giardia; ds.

XX OS Unidentified.
XX
XX CN1368384-A.
XX
XX 11-SEP-2002.
XX
XX 08-FEB-2001; 2001CN-00102614.
XX
XX 08-FEB-2001; 2001CN-00102614.
XX
XX (DIWE-) DIWEIHUAYU BIO TECHNOLOGY CO LTD BEIJING.
XX
XX Yu Y, Li H;
XX
XX WPI; 2003-854662/80.
XX
XX Genetically engineered vaccine of MUC-1 antigen for human breast cancer.
XX
XX Example 2; Page 4 (Disclosure); 14pp; Chinese.
XX
XX The invention relates to a method of preparation of a genetically
XX engineered vaccine for preventing and treating human breast cancer. The
XX method comprises fusing the coding gene of the Mycobacterium bovis heat
XX shock protein 65 (BCG HSP65) with the cytotoxic T lymphocyte (CTL)
XX epitope gene of MUC1 (antigen cell expressed by human breast cancer
XX cells) and the HSP65-MUC1 antigen CTL epitope is expressed in Escherichia
XX coli cells. This sequence represents DNA used in the method of the
XX invention.
XX
XX Sequence 156 BP; 37 A; 42 C; 58 G; 19 T; 0 U; 0 Other;
XX
XX
XX Query Match 47.7%; Score 28.6; DB 10; Length 156;
XX Best Local Similarity 67.8%; Pred. No. 2.4;
XX Matches 40; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
XX
XX 1 GGTAGTACAGCGCCGACATGAGCGCTCAGAGCGCCGAGTACGAGACCGGCGCC 59
XX 132 GGTTCACCGCTCCGCGGCTCAGGCTTACTCTGCTCCGACACCCGCTCGGCTCC 74
XX
XX
XX RESULT 14
XX ADK68629
XX ID ADK68629 standard; cDNA; 1800 BP.
XX
XX AC ADK68629;
XX
XX DT 06-MAY-2004 (first entry)
XX
XX DE HSP65-MUC1 antigen CTL epitope related cDNA #1.
XX
XX KM Human breast cancer; Mycobacterium bovis; heat shock protein 65; BCG;
XX HSP65; cytotoxic T lymphocyte; CTL; MUC1; HSP65-MUC1 antigen CTL epitope;
XX gene; ss.
XX
XX OS Unidentified.
XX
XX CN1368384-A.
XX
XX 11-SEP-2002.
XX
XX PD 08-FEB-2001; 2001CN-00102614.
XX
XX PR 08-FEB-2001; 2001CN-00102614.
XX
XX PA (DIWE-) DIWEIHUAYU BIO TECHNOLOGY CO LTD BEIJING.
XX
XX Yu Y, Li H;
XX
XX WPI; 2003-854662/80.
XX
XX P-PSDB; ADK68630.
XX
XX Genetically engineered vaccine of MUC-1 antigen for human breast cancer.

XX PS Claim 1; SEQ ID NO 1; 14pp; Chinese.
XX
XX CC The invention relates to a method of preparation of a genetically
XX engineered vaccine for preventing and treating human breast cancer. The
XX method comprises fusing the coding gene of the Mycobacterium bovis heat
XX shock protein 65 (BCG HSP65) with the cytotoxic T lymphocyte (CTL)
XX epitope gene of MUC1 (antigen cell expressed by human breast cancer
XX cells) and the HSP65-MUC1 antigen CTL epitope is expressed in Escherichia
XX coli cells. This sequence represents DNA used in the method of the
XX invention.
XX
XX Sequence 1800 BP; 347 A; 561 C; 616 G; 276 T; 0 U; 0 Other;
XX
XX
XX Query Match 47.7%; Score 28.6; DB 10; Length 1800;
XX Best Local Similarity 67.8%; Pred. No. 2.8;
XX Matches 40; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
XX
XX 1 GGTAGTACAGCGCCGACATGAGCGCTCAGAGCGCCGAGTACGAGACCGGCGCC 59
XX 1639 GGTTCACCGCTCCGCGGCTCAGGCTTACTCTGCTCCGACACCCGCTCGGCTCC 1697
XX
XX
XX RESULT 15
XX AD157666
XX ID AD157666 standard; cDNA; 4144 BP.
XX
XX AC AD157666;
XX
XX DT 22-APR-2004 (first entry)
XX
XX DE Human breast specific nucleic acid (BSNA) #37.
XX
XX KM Human; breast specific nucleic acid; BSNA; gene; ss; metatasis;
XX breast cancer; cytosstatic.
XX
XX OS Homo sapiens.
XX
XX PN WO2003106648-A2.
XX
XX PD 24-DEC-2003.
XX
XX PF 16-JUN-2003; 2003WO-US018934.
XX
XX PR 14-JUN-2002; 2002US-0389327P.
XX
XX PA (DIAD-) DIADEXUS INC.
XX
XX PI Salceda S, Macina RA, Turner LR, Sun Y, Liu C;
XX
XX DR WPI; 2004-082185/08.
XX
XX DR P-PSDB; AD157740.
XX
XX PT Novel isolated polypeptide comprising breast specific protein sequences,
XX useful for diagnosing or monitoring presence and metastases of breast
XX cancer in patient.
XX
XX PS Claim 1; SEQ ID NO 37; 370pp; English.
XX
XX CC The invention relates to human breast specific nucleic acids (BSNA) and
XX the breast specific proteins (BSP) they encode. The nucleic acids are
XX useful for determining the presence of a BSNA in a sample which involves
XX contacting the sample with a BSNA under conditions in which the BSNA will
XX selectively hybridize to a BSNA in the sample, and detecting the
XX hybridisation. The nucleic acids are useful for determining the presence
XX of a BSP in a sample which involves contacting the sample with suitable
XX reagent under conditions in which the reagent will selectively interact
XX with the BSP, and detecting the interaction of the reagent with a BSP in
XX the sample. The nucleic acids and proteins are useful for diagnosing or
XX monitoring the presence and metastases of breast cancer in a patient,
XX which involves determining an amount of nucleic acid or protein and
XX comparing the determined amount of nucleic acid or protein in a sample
XX of the patient to the amount of a breast specific marker in a normal

CC control, where a difference in the determined amount in the sample
 CC compared to the amount in the control is associated with the presence of
 CC breast cancer. The sequences are useful for treating a patient with
 CC breast cancer, involving administering a composition consisting of a BSNA
 CC or a BSP to a patient, where the administration induces an immune
 CC response against the breast cancer cell expressing the BSNA or BSP. This
 CC sequence represents a human BSNA of the invention.

XX
 SQ Sequence 4144 BP; 634 A; 1896 C; 1058 G; 552 T; 0 U; 4 Other;

Query Match 47.7%; Score 28.6; DB 12; Length 4144;
 Best Local Similarity 69.8%; Pred. No. 3;
 Matches 37; Conservative 1; Mismatches 15; Indels 0; Gaps 0;

Oy 7 ACAGGCCGACCCGCACTGCGCTCACGAGCCGCTCCGATACGAGACCGGCGCC 59
 |||||
 Db 699 ACCGCCGCCGCCGACGAYGTCTACCTCGGCCGCCGACACGAGCCGCGCC 751

Search completed: January 15, 2005, 20:36:12
 Job time : 171.5 secs

Query Match 100.0%; Score 60; DB 6; Length 2297;
 Best Local Similarity 100.0%; Pred. No. 6.1e-08;
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 GGTAGTACAGCGCCGACCATGGCGTCCGATGCCGATCGAGACCGCGCCT 60
 886 GGTAGTACAGCGCCGACCATGGCGTCCGATGCCGATCGAGACCGCGCCT 945

RESULT 2
 LOCUS BD225141 525 bp DNA linear PAT 17-JUL-2003
 DEFINITION Medicinal composition having antitumor effect and containing DNA
 encoding antigenic protein.
 ACCESSION BD225141
 VERSION BD225141.1 GI:33034911
 KEYWORDS JP 2002528519-A/2.
 SOURCE JP 2002528519-A/2.
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 Pallette, D., Massimo, A.M.D. and Desantis, R.
 Medicinal composition having antitumor effect and containing DNA
 encoding antigenic protein
 Patent: JP 2002528519-A 2 03-SEP-2002;
 MENARINI RICERCHE SPA
 OS Homo sapiens (human)
 PN JP 2002528519-A/2
 PD 03-SEP-2002
 PE 18-OCT-1999 JP 2000579265
 PR 30-OCT-1998 IT MI98A002330
 PI DINO PALLENTE, ANNA MARIA D MASSIMO, RITA DESANTIS PC
 A61K38/00, A61K35/76, A61K39/00, A61K48/00, A61P35/00, C12N15/09, PC
 A61K37/02,
 PC C12N15/00
 CC Medicinal composition having antitumor effect and containing
 CC antigenic protein
 CC DNA encoding
 CC key
 FH key
 FT source
 FT Location/Qualifiers
 1..525
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

FEATURES
 source

ORIGIN

Query Match 61.0%; Score 36.6; DB 6; Length 525;
 Best Local Similarity 76.3%; Pred. No. 0.74; Indels 14; Gaps 0;
 Matches 45; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Db 1 GGTAGTACAGCGCCGACCATGGCGTCCGATGCCGATCGAGACCGCGCCT 59
 256 GGTAGTACAGCGCCGACCATGGCGTCCGATGCCGATCGAGACCGCGCCT 314

RESULT 3
 LOCUS BD225147 891 bp DNA linear PAT 17-JUL-2003
 DEFINITION Medicinal composition having antitumor effect and containing DNA
 encoding antigenic protein.
 ACCESSION BD225147
 VERSION BD225147.1 GI:33034917
 KEYWORDS JP 2002528519-A/8.
 SOURCE JP 2002528519-A/8.
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 Pallette, D., Massimo, A.M.D. and Desantis, R.

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

FEATURES
 source

TITLE
 JOURNAL
 COMMENT

Db 1 GGTAGTACAGCGCCGACCATGGCGTCCGATGCCGATCGAGACCGCGCCT 59
 622 GGTAGTACAGCGCCGACCATGGCGTCCGATGCCGATCGAGACCGCGCCT 680

RESULT 4
 LOCUS BD225144 1371 bp DNA linear PAT 17-JUL-2003
 DEFINITION Medicinal composition having antitumor effect and containing DNA
 encoding antigenic protein.
 ACCESSION BD225144
 VERSION BD225144.1 GI:33034914
 KEYWORDS JP 2002528519-A/5.
 SOURCE JP 2002528519-A/5.
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 Pallette, D., Massimo, A.M.D. and Desantis, R.
 Medicinal composition having antitumor effect and containing DNA
 encoding antigenic protein
 Patent: JP 2002528519-A 5 03-SEP-2002;
 MENARINI RICERCHE SPA
 OS Homo sapiens (human)
 PN JP 2002528519-A/5
 PD 03-SEP-2002
 PE 18-OCT-1999 JP 2000579265
 PR 30-OCT-1998 IT MI98A002330
 PI DINO PALLENTE, ANNA MARIA D MASSIMO, RITA DESANTIS PC
 A61K38/00, A61K35/76, A61K39/00, A61K48/00, A61P35/00, C12N15/09, PC
 A61K37/02,
 PC C12N15/00
 CC Medicinal composition having antitumor effect and containing
 CC antigenic protein
 CC DNA encoding
 CC key
 FH key
 FT source
 FT Location/Qualifiers
 1..1371
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

FEATURES
 source

ORIGIN

Query Match 61.0%; Score 36.6; DB 6; Length 1371;
Best Local Similarity 76.3%; Pred. No. 0.68;
Matches 45; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 1 GGTAGTACAGCCGACCCGACATGCGCTCAGAGCGCTCCGATACGAGACCGGCGCC 59
Db 325 GGAGTACCGCTCCACACGACGCTTACTCGGCTCCGATACGAGCGCGCC 383

RESULT 5
BD225150 1737 bp DNA linear PAT 17-JUL-2003
LOCUS BD225150
DEFINITION Medicinal composition having antitumor effect and containing DNA
encoding antigenic protein.
ACCESSION BD225150
VERSION BD225150.1 GI:33034920
KEYWORDS JP 2002528519-A/11.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
TITLE 1 (bases 1 to 1737)
JOURNAL Pallen, D., Massimo, A.M.D. and Desantis, R.
Medicinal composition having antitumor effect and containing DNA
encoding antigenic protein
Patent: JP 2002528519-A 11 03-SEP-2002;
MENARINI RICERCHE SPA
OS Homo sapiens (human)
PD JP 2002528519-A/11
PN 03-SEP-2002 JP 2000579265
PF 18-OCT-1999 JP 2000579265
PR 30-OCT-1998 IT MI98A002330
PI DINO PALLENTI, ANNA MARIA D MASSIMO, RITA DESANTIS PC
A61K38/00, A61K35/76, A61K39/00, A61K48/00, A61P35/00, C12N15/09, PC
A61K37/02,
PC C12N15/00
CC Medicinal composition having antitumor effect and containing
CC DNA encoding
CC antigenic protein
FH Key Location/Qualifiers
FT source 1.1737
FT Location/Qualifiers
1.1737
location/Qualifiers
1.1737
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 61.0%; Score 36.6; DB 6; Length 1371;
Best Local Similarity 76.3%; Pred. No. 0.67;
Matches 45; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 1 GGTAGTACAGCCGACCCGACATGCGCTCAGAGCGCTCCGATACGAGACCGGCGCC 59
Db 691 GGAGTACCGCTCCACACGACGCTTACTCGGCTCCGATACGAGCGCGCC 749

RESULT 6
AX192396/c 120 bp DNA linear PAT 15-AUG-2001
LOCUS AX192396
DEFINITION Sequence 2 from Patent EP1103623.
ACCESSION AX192396
VERSION AX192396.1 GI:15210363
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
TITLE 1
JOURNAL Taylor-Papadimitriou, J., Burchell, J. and Gendler, S.

TITLE Human mucin core protein: nucleic acid probes, peptide fragments and antibodies thereto, and uses thereof in diagnostic and therapeutic methods

JOURNAL Patent: EP 1103623-A 2 30-MAY-2001;
IMPERIAL CANCER RESEARCH TECHNOLOGY LIMITED (GB)

FEATURES
source 1.120
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 45.7%; Score 27.4; DB 6; Length 120;
Best Local Similarity 69.8%; Pred. No. 4.9e+02;
Matches 37; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 7 ACAGGCGCACCCGACATGCGCTCAGAGCGCTCCGATACGAGACCGGCGCC 59
Db 81 ACCGCGCGCGCGCGCGCGCGCTCAGAGCGCGCGCGCGCGCGCGCGCGCGCC 29

RESULT 7
BD000571 120 bp DNA linear PAT 31-JAN-2002
LOCUS BD000571/c
DEFINITION Human polymorphic epithelial mucin core protein and nucleic acid
encoding the protein.
ACCESSION BD000571
VERSION BD000571.1 GI:18623684
KEYWORDS JP 2000333675-A/2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
TITLE 1 (bases 1 to 120)
JOURNAL Papadimitriou, J.T., Jendora, S. and Bachieri, J.
Human polymorphic epithelial mucin core protein and nucleic acid
encoding the protein
Patent: JP 2000333675-A 2 05-DEC-2000;
IMPERIAL CANCER RESEARCH TECHNOLOGY LTD
OS Homo sapiens (human)
PD JP 2000333675-A/2
PN 05-DEC-2000 JP 2000125724
PF 26-APR-2000 JP 2000125724
PR 07-JAN-1987 GB 8700269 07-JAN-1987 GB 8700279 PR
22-APR-1987 US 041306, 09-NOV-1987 GB 8726172 PI JOICE
TAYLOR PAPADIMITRIOU, SANDRA JENDORA, JOY BACHIERU PC
C12N15/02, A61K38/00, A61K39/395, A61K49/00, A61P35/00, PC
C07K14/47,
PC C07K16/44, C12N5/10, C12P21/08// (C12N15/02, C12R1:91), (C12N5/10,
PC C12R1:91),
PC C12N15/00, A61K37/02, C12N5/00, (C12N15/00, C12R1:91), (C12N5/00,
PC C12R1:91)
CC
FH Key Location/Qualifiers
FT source 1.120
FT Location/Qualifiers
1.120
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 45.7%; Score 27.4; DB 6; Length 120;
Best Local Similarity 69.8%; Pred. No. 4.9e+02;
Matches 37; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 7 ACAGGCGCACCCGACATGCGCTCAGAGCGCTCCGATACGAGACCGGCGCC 59
Db 81 ACCGCGCGCGCGCGCGCGCGCTCAGAGCGCGCGCGCGCGCGCGCGCGCGCC 29

RESULT 8

E08763 LOCUS E08763 180 bp RNA linear PAT 29-SEP-1997
 DEFINITION cDNA encoding N-terminal fragment of human glycoprotein 39.
 ACCESSION E08763
 VERSION E08763.1 GI:2176875
 KEYWORDS JP 1995051065-A/1.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 1 (bases 1 to 180)
 Masuzawa,Y., Muramatsu,T. and Miyachi,T.
 GLYCOPROTEIN 39 GENE
 Patent: JP 1995051065-A 1 28-FEB-1995.
 NIPPON KOUTAI KENKYUSHO:KK, KAGOSHIMA UNIV
 OS Homo sapiens (human)
 PN JP 1995051065-A/1
 PD 28-FEB-1995
 PF 21-FEB-1992 JP 1992035085
 PI MASUZAWA YASUSHI, MURAMATSU TAKASHI, MIYACHI TERUO PC
 C12N15/09,G01N33/50//C07K14/00,C12P21/02;
 CC strandedness: Double;
 CC topology: Linear;
 CC hypothetical: No;
 CC anti-sense: No;
 CC Feature is identified by similarity;
 FH Key Location/Qualifiers
 FT source 1..180
 FT /organism="Homo sapiens"
 FT /cell_type="gastric cancer"
 FT /cell_line="KATO-3"
 FT /clone_11b="lambda gpi1 KATO-3 cDNA library"
 FT /clone="lambda KP39"
 FT mat_peptide 1..180
 FT /product="N-terminal fragment of human FT
 FT glycoprotein 39"
 FT repeat_region 1..180
 FT repeat_unit 1..60.
 FT Location/Qualifiers
 1..180
 /organism="Homo sapiens"
 /mol_type="genomic RNA"
 /db_xref="taxon:9606"
 ORIGIN
 Query Match 45.7%; Score 27.4; DB 6; Length 180;
 Best Local Similarity 69.8%; Pred. No. 4.7e+02;
 Matches 37; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
 Oy 7 ACAAGCGCCACCCGACATGCGTACAGAGCGCTCCGATACGAGACCGCGCC 59
 Db 127 ACCGCGCCGCGACGCGTGTCACTCGGCCCGACACGAGCGCGCC 179
 RESULT 9
 AF423031 1414 bp mRNA linear SYN 10-JUL-2003
 LOCUS AF423031
 DEFINITION Synthetic construct Homo sapiens mucin variant MUC1-CT58 (MUC1)
 ACCESSION AF423031
 VERSION AF423031.1 GI:19338621
 KEYWORDS
 SOURCE synthetic construct
 ORGANISM synthetic construct
 1 (bases 1 to 1414)
 Hinojosa-Kurtzberg,A.M., Johansson,M.E., Madsen,C.S., Hansson,G.C.
 and Gendler,S.J.
 Novel MUC1 splice variants contribute to mucin overexpression in
 CTR-deficient mice
 Am. J. Physiol. Gastrointest. Liver Physiol. 284 (5), G853-G862

(2003)
 MEDLINE 22570517
 PUBMED 12529261
 REFERENCE 2 (bases 1 to 1414)
 AUTHORS Hinojosa-Kurtzberg,A.M. and Gendler,S.J.
 TITLE Direct Submission
 JOURNAL Submitted (21-SEP-2001) Biochemistry, Mayo Clinic, 13400 E. Shea
 Boulevard, Scottsdale, AZ 85259, USA
 FEATURES
 source Location/Qualifiers
 1..1414
 /organism="synthetic construct"
 /mol_type="mRNA"
 /db_xref="taxon:32630"
 /note="Homo sapiens gene in transgenic Mus musculus
 C57BL/6; isolated from intestinal mucosa"
 1..1414
 /gene="MUC1"
 1..1386
 /gene="MUC1"
 /note="alternatively spliced; contains exon 6b resulting
 in variant carboxy-terminal domain; lacks sites for
 beta-catenin and Grb2 interactions; derived from Homo
 sapiens"
 /codon_start=1
 /transl_table=1
 /product="mucin variant MUC1-CT58"
 /protein_id="AAU86735.1"
 /db_xref="GI:19338622"
 /translation="MPTGTSPFFLLLLTLVLTVTGSGHASTPGKEKTSATORSS
 VPSSTENKAVSWTSVLSHSFGSGSTTQGDVTLAPATPESGSAATWGDVSV
 VTRPALSTTPPAHDVTSAPDNKAPAGSTPAHGVTSAPDRPAPAGHGVTS
 APDRPPLAGSTAPVHVNTVSASGASGASATVHNGTSATTPASKSPSTSHH
 SDPTTLASTSTVDASTHSTVPLTSSNSTSPSTSTGVSFFLSHISNLPNS
 SLEPSTDYQTELRDISMPLQIVKGGFLGKSNKRPSSVVQTLAEFGTIN
 HDVETPNQVQTEALRSRYNLTSDVSVSPFSSQSGAGVGLALVIVCLVA
 LAIVLIALAVCCCRKNYQGLDIFPARDTHPMSELYPTHTGRVPPSSYDRSPE
 KEMRVDRERLA"
 ORIGIN
 Query Match 45.7%; Score 27.4; DB 12; Length 1414;
 Best Local Similarity 69.8%; Pred. No. 4e+02;
 Matches 37; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
 Oy 7 ACAAGCGCCACCCGACATGCGTACAGAGCGCTCCGATACGAGACCGCGCC 59
 Db 391 ACCGCGCCGCGACGCGTGTCACTCGGCCCGACACGAGCGCGCC 443
 RESULT 10
 CO715242 1455 bp DNA linear PAT 03-FEB-2004
 LOCUS CO715242
 DEFINITION Sequence 1176 from Patent WO02068579.
 ACCESSION CO715242
 VERSION CO715242.1 GI:42276099
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1
 AUTHORS Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
 TITLE Kits, such as nucleic acid arrays, comprising a majority of
 humanexons or transcripts, for detecting expression and other uses
 thereof
 JOURNAL Patent: WO 02068579-A 1176 06-SEP-2002;
 PE Corporation (NY) (US)
 FEATURES
 source Location/Qualifiers
 1..1455
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 ORIGIN

Query Match	45.7%	Score 27.4	DB 6	Length 1457
Best Local Similarity	69.8%	Pred. No. 4e+02		
Matches	37	Conservative	0	Mismatches 16; Indels 0; Gaps 0
RESULT 11				
LOCUS	AX959914	1457 bp	DNA	linear
DEFINITION	Sequence 19 from Patent WO03100060.			PAT 14-JAN-2004
ACCESSION	AX959914			
VERSION	AX959914.1	GI:40880143		
KEYWORDS				
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
AUTHORS	1			
TITLE	Burden,N.G., Ellis,J.H. and Hamblin,P.A.			
JOURNAL	Muc-1 antigen with reduced number of vnter repeat units			
FEATURES	Patent: WO 03100060-A 19 04-DEC-2003;			
source	GLAXO GROUP LIMITED (GB)			
Location/Qualifiers				
1..1457				
/organism="Homo sapiens"				
/mol_type="unassigned DNA"				
/db_xref="taxon:9606"				
ORIGIN				
Query Match	45.7%	Score 27.4	DB 6	Length 1457
Best Local Similarity	69.8%	Pred. No. 4e+02		
Matches	37	Conservative	0	Mismatches 16; Indels 0; Gaps 0
Db	409	ACCGCCCCCAGCCACGCGTGTCTCCTCGGCCCCGACACACGCGGCCCC	461	
Qy	7	ACAGGCGCACCCGCATCGCGCTCACGAGCGCTCGGATACGACACGCGGCC	59	
LOCUS	AX093798	1572 bp	DNA	linear
DEFINITION	Sequence 1 from Patent WO0118035.			PAT 30-MAR-2001
ACCESSION	AX093798			
VERSION	AX093798.1	GI:13510057		
KEYWORDS				
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
AUTHORS	1			
TITLE	Taylor-Papadimitriou,J., Heukamp,L.C., Offringa,R., Melief,C.J., Acres,B. and Thomas,M.			
JOURNAL	Muc-1 derived peptides			
FEATURES	Patent: WO 0118035-A 1 15-MAR-2001;			
source	TRANSGENE S.A. (FR) ; IMPERIAL CANCER RESEARCH TECHNOLOGY LIMITED (GB)			
Location/Qualifiers				
1..1572				
/organism="Homo sapiens"				
/mol_type="unassigned DNA"				
/db_xref="taxon:9606"				
58..1545				
/note="unamed protein product"				
/codon_start=1				
/protein_id="CAC35422.1"				
/db_xref="GI:13510058"				
/translation="MTPTGTGPPFLILLTLVLTVTGSGHASTRGGEKTSATORS				
VPSSTKXAVMTSSVLSHSGSGSSTGQDVTLAATPAPASGAATGQDVTSP				
VTPRALGTTTPADVTSAIPDKRAPGSTAPAPAGVTSAIPTRPPGSTAPALAGVTS				
APDTRPAGSTAPPAHGVTSAPDNRPALGSTAPAPVHNTASGASGSASTLVHNGTS				

Query Match	45.7%;	Score 27.4;	DB 6;	Length 1572;
Best Local Similarity	69.8%;	Pred. No. 3.9e+02;		
Matches 37;	Conservative 0;	Mismatches 16;	Indels 0;	Gaps 0;
Db	508	ACGCGCCCGCCGCGCCACGGGTGTCACTCGGCCCCGAGACACGAGCGCGCCCC	560	
RESULT 13				
LOCUS	CQ771290	1721 bp	DNA	linear PAT 04-MAR-2004
DEFINITION	Sequence 1 from Patent WO2004009632.			
ACCESSION	CQ771290			
VERSION	CQ771290.1	GI:45125397		
KEYWORDS				
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
AUTHORS	Goletz,S. and Karsten,U.			
TITLE	Method for the production of an immunostimulating mucin (muc1)			
JOURNAL	Patent: WO 2004009632-A 1 29-JAN-2004;			
FEATURES	Nemod Immunotherapie AG (DE)			
source	Location/Qualifiers			
ORIGIN	1..1721			
	/organism="Homo sapiens"			
	/mol_type="unassigned DNA"			
	/db_xref="taxon:9606"			
Query Match	45.7%;	Score 27.4;	DB 6;	Length 1721;
Best Local Similarity	69.8%;	Pred. No. 3.9e+02;		
Matches 37;	Conservative 0;	Mismatches 16;	Indels 0;	Gaps 0;
Db	568	ACCGCGCCCGCAGCCCGGTGTCACTCGGCCCCGAGACACGAGCGCGCCC	620	
RESULT 14				
LOCUS	AR492306	1721 bp	DNA	linear PAT 15-MAY-2004
DEFINITION	Sequence 3 from patent US 6716627.			
ACCESSION	AR492306			
VERSION	AR492306.1	GI:47260880		
KEYWORDS				
SOURCE	Unknown.			
ORGANISM	Unknown.			
REFERENCE	Unclassified.			
AUTHORS	1 (bases 1 to 1721)			
TITLE	Dobie,K.W.			
JOURNAL	Antisense modulation of mucin 1, transmembrane expression			
FEATURES	Patent: US 6716627-A 3 06-APR-2004;			
source	Location/Qualifiers			
	1..1721			
	/organism="unknown"			
	/mol_type="genomic DNA"			
ORIGIN				
Query Match	45.7%;	Score 27.4;	DB 6;	Length 1721;
Best Local Similarity	69.8%;	Pred. No. 3.9e+02;		
Matches 37;	Conservative 0;	Mismatches 16;	Indels 0;	Gaps 0;
Db	7	ACAGCGCACCCCGACATGGCGGTCAAGCGCTCCGATACGAGACCGCGCCC	59	
ORIGIN				

Db 568 ACCGCGCCCGCAGCGGTGTCACCTCGGCCCGAGACACCGCGGCCCC 620

RESULT 15

AX335860

1721 bp DNA

linear

PAT 09-JAN-2002

Sequence 6369 from Patent WO0194629.

ACCESSION

AX335860

VERSION

AX335860.1 GI:18126579

KEYWORDS

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Young, P.E., Augustus, M., Carter, K.C., Edner, R., Endress, G.,

Horrigan, S., Soppet, D.R. and Weaver, Z.

Cancer gene determination and therapeutic screening using signature

gene sets

Patent: WO 0194629-A 6369 13-DEC-2001;

Avalon Pharmaceuticals (US)

FEATURES

location/Qualifiers

1..1721

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

ORIGIN

Query Match

Best Local Similarity 45.7%; Score 27.4; DB 6; Length 1721;

Matches 37; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 7 ACAGCGCCACCCGACATGCGCTCAGCGCGCTCCGGATACGAGACCGGCGCC 59

Db 568 ACCGCGCCCGCAGCGGTGTCACCTCGGCCCGAGACACCGCGGCCCC 620

Search completed: January 15, 2005, 22:36:55
Job time : 723.8 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2005, 22:37:00 : Search time 183.4 Seconds
(without alignments)
1879.789 Million cell updates/sec

Title: US-10-057-136-11

Perfect score: 60
Sequence: 1 gggtcaacgcagcaccctccagc.....ccgacacccgtccagctccg 60

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4300275 seqs, 2872944193 residues

Total number of hits satisfying chosen parameters: 8600550

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
21: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	60	100.0	60	US-10-057-136-11	Sequence 11, Appl
2	60	100.0	2287	US-10-406-317-41	Sequence 41, Appl
3	38.2	63.7	60	US-10-057-136-5	Sequence 5, Appl
4	36.6	61.0	60	US-10-057-136-6	Sequence 6, Appl
5	36	60.0	60	US-10-057-136-2	Sequence 2, Appl
6	36	60.0	78	US-10-057-136-13	Sequence 13, Appl
7	36	60.0	1424	US-10-447-839A-75	Sequence 75, Appl
8	36	60.0	1428	US-10-447-839A-20	Sequence 20, Appl
9	36	60.0	1527	US-10-057-136-19	Sequence 19, Appl
10	36	60.0	1799	US-10-447-839A-19	Sequence 19, Appl
11	36	60.0	1804	US-09-964-824A-573	Sequence 573, Appl
12	36	60.0	1804	US-10-029-517-17	Sequence 17, Appl

13	36	60.0	1804	US-10-717-597-30	Sequence 30, Appl
14	36	60.0	1804	US-10-775-920-84	Sequence 84, Appl
15	36	60.0	1823	US-10-101-510-139	Sequence 339, Appl
16	36	60.0	4139	US-09-964-824A-105	Sequence 105, Appl
17	36	60.0	4139	US-09-964-824A-578	Sequence 578, Appl
18	36	60.0	4139	US-09-864-864-334	Sequence 334, Appl
19	36	60.0	4139	US-09-880-107-2121	Sequence 2121, Appl
20	36	60.0	4139	US-09-968-007A-751	Sequence 751, Appl
21	36	60.0	4139	US-10-171-311-157	Sequence 157, Appl
22	36	60.0	4139	US-10-177-293-310	Sequence 310, Appl
23	36	60.0	4139	US-10-440-464-155	Sequence 155, Appl
24	36	60.0	4139	US-10-734-564-53	Sequence 53, Appl
25	36	60.0	4139	US-10-775-920-80	Sequence 80, Appl
26	36	60.0	4139	US-10-775-920-85	Sequence 85, Appl
27	36	60.0	8181	US-10-447-839A-18	Sequence 18, Appl
28	36	60.0	8186	US-10-247-703-23	Sequence 23, Appl
29	36	60.0	8186	US-10-029-517-19	Sequence 19, Appl
30	34.4	57.3	572	US-10-029-517-18	Sequence 18, Appl
31	34.4	57.3	572	US-10-775-920-93	Sequence 93, Appl
32	34.4	57.3	1721	US-09-864-864-280	Sequence 280, Appl
33	34.4	57.3	1721	US-09-967-768A-224	Sequence 224, Appl
34	34.4	57.3	1721	US-10-247-703-21	Sequence 21, Appl
35	34.4	57.3	1721	US-10-097-340-211	Sequence 211, Appl
36	34.4	57.3	1721	US-10-171-311-155	Sequence 155, Appl
37	34.4	57.3	1721	US-10-007-926A-58	Sequence 58, Appl
38	34.4	57.3	1721	US-10-029-517-3	Sequence 3, Appl
39	34.4	57.3	1721	US-10-172-118-775	Sequence 775, Appl
40	34.4	57.3	1721	US-10-342-887-775	Sequence 775, Appl
41	34.4	57.3	1721	US-10-775-920-88	Sequence 88, Appl
42	34.4	57.3	2026	US-10-198-846-12589	Sequence 12589, A
43	34.4	57.3	2238	US-10-775-920-87	Sequence 87, Appl
44	34.4	57.3	2678	US-10-252-157-103	Sequence 103, Appl
45	33.6	56.0	518	US-10-247-703-38	Sequence 38, Appl

ALIGNMENTS

RESULT 1
US-10-057-136-11
Sequence 11, Application US/10057336
Publication No. US20030021770A1
GENERAL INFORMATION:
APPLICANT: SCHLON, JEFFREY
APPLICANT: KANTOR, JUDITH
APPLICANT: KUFE, DONALD
APPLICANT: PANICALI, DENNIS
APPLICANT: GRITZ, LINDA
TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
TITLE OF INVENTION: TUMOR-ASSOCIATED ANTIGEN
FILE REFERENCE: 700953/47113C
CURRENT APPLICATION NUMBER: US/10/057,136
CURRENT FILING DATE: 2002-01-25
PRIOR APPLICATION NUMBER: 09/366,670
PRIOR FILING DATE: 1999-08-03
PRIOR APPLICATION NUMBER: PCT/US98/03693
PRIOR FILING DATE: 1998-02-24
PRIOR APPLICATION NUMBER: 60/038,253
PRIOR FILING DATE: 1997-02-24
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 11
LENGTH: 60
TYPE: DNA
ORGANISM: Homo sapiens
US-10-057-136-11

Query Match 100.0%; Score 60; DB 14; Length 60;
Best Local Similarity 100.0%; Pred. No. 1.2e-11;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 GGTTCAACGCAGCCTCCAGACAGGAGTGCAGTCCGACCGACCGCTCCAGCTCCG 60
|||||

Db 1 GGTTCACGGCAGCTCCAGCAGCAGAGTCACTTGACCCGACACCCGTCAGCTCCG 60

RESULT 2

US-10-406-317-41
; Sequence 41, Application US/10406317
; Publication No. US20040019195A1
; GENERAL INFORMATION:
; APPLICANT: Schiom, Jeffrey;
; APPLICANT: Hodge, James;
; APPLICANT: Panicali, Dennis
; TITLE OF INVENTION: A recombinant vector expressing multiple constitutively
; FILE REFERENCE: 38163-0189
; CURRENT APPLICATION NUMBER: US/10/406,317
; CURRENT FILING DATE: 2003-04-04
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: PCT/US99/26866
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/111,582
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41
; LENGTH: 2297
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: VECTOR
US-10-406-317-41

Query Match 100.0%; Score 60; DB 16; Length 2297;
Best Local Similarity 100.0%; Pred. No. 1,3e-11;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCACGGCAGCTCCAGCAGCAGAGTCACTTGACCCGACACCCGTCAGCTCCG 60
Db 826 GGTTCACGGCAGCTCCAGCAGCAGAGTCACTTGACCCGACACCCGTCAGCTCCG 885

RESULT 3

US-10-057-136-5
; Sequence 5, Application US/10057136
; Publication No. US20030021770A1
; GENERAL INFORMATION:
; APPLICANT: SCHIOM, JEFFREY
; APPLICANT: KANTOR, JUDITH
; APPLICANT: KUFE, DONALD
; APPLICANT: PANICALI, DENNIS
; APPLICANT: GRITZ, LINDA
; TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
; FILE REFERENCE: 700953/47113C
; CURRENT APPLICATION NUMBER: US/10/057,136
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/366,670
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: PCT/US98/03693
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/038,253
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-136-5

Query Match 63.7%; Score 38.2; DB 14; Length 60;
Best Local Similarity 78.0%; Pred. No. 0.00059;

Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 GGTTCACGGCAGCTCCAGCAGCAGAGTCACTTGACCCGACACCCGTCAGCTCCG 59
Db 1 GGATCCACGGCGCGCGCTCGCAGCAGAGTCACTTGCGGCGCCGACACCGCGCGCTCC 59

RESULT 4

US-10-057-136-6
; Sequence 6, Application US/10057136
; Publication No. US20030021770A1
; GENERAL INFORMATION:
; APPLICANT: SCHIOM, JEFFREY
; APPLICANT: KANTOR, JUDITH
; APPLICANT: KUFE, DONALD
; APPLICANT: PANICALI, DENNIS
; APPLICANT: GRITZ, LINDA
; TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
; FILE REFERENCE: 700953/47113C
; CURRENT APPLICATION NUMBER: US/10/057,136
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/366,670
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: PCT/US98/03693
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/038,253
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-136-6

Query Match 61.0%; Score 36.6; DB 14; Length 60;
Best Local Similarity 76.3%; Pred. No. 0.0021;
Matches 45; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 GGTTCACGGCAGCTCCAGCAGCAGAGTCACTTGACCCGACACCCGTCAGCTCCG 59
Db 1 GGTTCACGGCAGCTCCAGCAGCAGAGTCACTTGCGGCGCGCTCCAGATATCGCCAGCTCC 59

RESULT 5

US-10-057-136-2
; Sequence 2, Application US/10057136
; Publication No. US20030021770A1
; GENERAL INFORMATION:
; APPLICANT: SCHIOM, JEFFREY
; APPLICANT: KANTOR, JUDITH
; APPLICANT: KUFE, DONALD
; APPLICANT: PANICALI, DENNIS
; APPLICANT: GRITZ, LINDA
; TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
; FILE REFERENCE: 700953/47113C
; CURRENT APPLICATION NUMBER: US/10/057,136
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/366,670
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: PCT/US98/03693
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/038,253
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-136-2

Query Match 60.0%; Score 36; DB 14; Length 60;
Best Local Similarity 75.0%; Pred. No. 0.0035;
Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 GGTTCACGGCAGCCTTCACGACGAGTCACGCTGACACCCGACCCGTCAGCTCCG 60
DB 1 GGCTCCACGGCCCCCCCCACGACGAGTGTACCTCGCCCCGACACGAGCGGCCCG 60

RESULT 6
US-10-057-136-13
Sequence 13, Application US/10057136
Publication No. US2003002170A1

GENERAL INFORMATION:
APPLICANT: SCHLOM, JEFFREY
APPLICANT: KANTOR, JUDITH
APPLICANT: KUFE, DONALD
APPLICANT: PANICALI, DENNIS
APPLICANT: GRITZ, LINDA
TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
FILE REFERENCE: 700953/47113C
CURRENT APPLICATION NUMBER: US/10/057,136

PRIOR FILING DATE: 2002-01-25
PRIOR APPLICATION NUMBER: 09/366,670
PRIOR FILING DATE: 1999-08-03
PRIOR APPLICATION NUMBER: PCT/US98/03693
PRIOR FILING DATE: 1998-02-24
PRIOR APPLICATION NUMBER: 60/038,253
PRIOR FILING DATE: 1997-02-24
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 13
LENGTH: 78
TYPE: DNA
ORGANISM: Homo sapiens
US-10-057-136-13

Query Match 60.0%; Score 36; DB 14; Length 78;
Best Local Similarity 75.0%; Pred. No. 0.0035;
Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 GGTTCACGGCAGCCTTCACGACGAGTCACGCTGACACCCGTCAGCTCCG 60
DB 1 GGCTCCACGGCAGCCTTCACGACGAGTGTACCTCGCCCCGACACGAGCGGCCCG 60

RESULT 7
US-10-447-839A-75/c
Sequence 75, Application US/10447839A
Publication No. US20040018181A1

GENERAL INFORMATION:
APPLICANT: KuFe, Donald W.
APPLICANT: Kharbanda, Surender
APPLICANT: Weitman, Steven D.
TITLE OF INVENTION: MUC1 INTERFERENCE RNA COMPOSITIONS AND METHODS DERIVED THEREFROM
FILE REFERENCE: 1000.05.009
CURRENT APPLICATION NUMBER: US/10/447,839A
PRIOR FILING DATE: 2003-05-29
PRIOR APPLICATION NUMBER: 10/293,391
PRIOR FILING DATE: 2002-11-13
PRIOR APPLICATION NUMBER: 09/951,938
PRIOR FILING DATE: 2001-09-11
PRIOR APPLICATION NUMBER: 60/231,841
PRIOR FILING DATE: 2000-09-11
NUMBER OF SEQ ID NOS: 109
SOFTWARE: PatentIn version 3.2
SEQ ID NO 75

LENGTH: 1424
TYPE: RNA
ORGANISM: ARTIFICIAL
FEATURE:

OTHER INFORMATION: Synthesized Sequence
US-10-447-839A-75

Query Match 60.0%; Score 36; DB 16; Length 1424;
Best Local Similarity 75.0%; Pred. No. 0.0036;
Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 GGTTCACGGCAGCCTTCACGACGAGTCACGCTGACACCCGACCCGTCAGCTCCG 60
DB 1041 GGCTCCACGGCCCCCCCCACGACGAGTGTACCTCGCCCCGACACGAGCGGCCCG 982

RESULT 8
US-10-447-839A-20
Sequence 20, Application US/10447839A
Publication No. US20040018181A1

GENERAL INFORMATION:
APPLICANT: KuFe, Donald W.
APPLICANT: Kharbanda, Surender
APPLICANT: Weitman, Steven D.
TITLE OF INVENTION: MUC1 INTERFERENCE RNA COMPOSITIONS AND METHODS DERIVED THEREFROM
FILE REFERENCE: 1000.05.009
CURRENT APPLICATION NUMBER: US/10/447,839A
PRIOR FILING DATE: 2003-05-29
PRIOR APPLICATION NUMBER: 10/293,391

PRIOR FILING DATE: 2002-11-13
PRIOR APPLICATION NUMBER: 09/951,938
PRIOR FILING DATE: 2001-09-11
PRIOR APPLICATION NUMBER: 60/231,841
PRIOR FILING DATE: 2000-09-11
NUMBER OF SEQ ID NOS: 109
SOFTWARE: PatentIn version 3.2

SEQ ID NO 20
LENGTH: 1428
TYPE: RNA
ORGANISM: RNA
US-10-447-839A-20

Query Match 60.0%; Score 36; DB 16; Length 1428;
Best Local Similarity 70.0%; Pred. No. 0.0036;
Matches 42; Conservative 3; Mismatches 15; Indels 0; Gaps 0;

QY 1 GGTTCACGGCAGCCTTCACGACGAGTCACGCTGACACCCGTCAGCTCCG 60
DB 385 GGCTCCACGGCCCCCCCCACGACGAGTGTACCTCGCCCCGACACGAGCGGCCCG 444

RESULT 9
US-10-057-136-19
Sequence 19, Application US/10057136
Publication No. US2003002170A1

GENERAL INFORMATION:
APPLICANT: SCHLOM, JEFFREY
APPLICANT: KANTOR, JUDITH
APPLICANT: KUFE, DONALD
APPLICANT: PANICALI, DENNIS
APPLICANT: GRITZ, LINDA
TITLE OF INVENTION: TUMOR-ASSOCIATED ANTIGEN
FILE REFERENCE: 700953/47113C
CURRENT APPLICATION NUMBER: US/10/057,136
PRIOR FILING DATE: 2002-01-25
PRIOR APPLICATION NUMBER: 09/366,670
PRIOR FILING DATE: 1999-08-03
PRIOR APPLICATION NUMBER: PCT/US98/03693
PRIOR FILING DATE: 1998-02-24
PRIOR APPLICATION NUMBER: 60/038,253
PRIOR FILING DATE: 1997-02-24
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 19
LENGTH: 1527
TYPE: DNA

ORGANISM: Homo sapiens
US-10-057-136-19

Query Match
Best Local Similarity 60.0%; Score 36; DB 14; Length 1527;
Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 GGTTCACGGCAGCTCCGACACGAGTCACTGTCACCCGACACCCGTCAGCTCCG 60
DB 226 GGCTCCACGGCCCCCCCCCAGCCAGGTGTCACTCGGCCCCGACACCAAGCGGCCCG 285

RESULT 10
US-10-447-839A-19

Sequence 19, Application US/10447839A
Publication No. US20040018181A1
GENERAL INFORMATION:
APPLICANT: Kufe, Donald W.
APPLICANT: Kharbanda, Sreender
APPLICANT: Weisman, Steven D.
TITLE OF INVENTION: MUC1 INTERFERENCE RNA COMPOSITIONS AND METHODS DERIVED THEREFROM
FILE REFERENCE: 1000.05.009
CURRENT APPLICATION NUMBER: US/10/447,839A
PRIOR FILING DATE: 2003-05-29
PRIOR APPLICATION NUMBER: 10/293,391
PRIOR FILING DATE: 2002-11-13
PRIOR APPLICATION NUMBER: 09/951,938
PRIOR FILING DATE: 2001-09-11
PRIOR APPLICATION NUMBER: 60/231,841
NUMBER OF SEQ ID NOS: 109
SOFTWARE: PatentIn version 3.2
SEQ ID NO 19
LENGTH: 1799
TYPE: RNA
ORGANISM: RNA
US-10-447-839A-19

Query Match
Best Local Similarity 70.0%; Score 36; DB 16; Length 1799;
Matches 42; Conservative 3; Mismatches 15; Indels 0; Gaps 0;

QY 1 GGTTCACGGCAGCTCCGACACGAGTCACTGTCACCCGACACCCGTCAGCTCCG 60
DB 453 GGCTCCACGGCCCCCCCCCAGCCAGGTGTCACTCGGCCCCGACACCAAGCGGCCCG 512

RESULT 11
US-09-964-824A-573

Sequence 573, Application US/09964824A
Patent No. US20020102531A1
GENERAL INFORMATION:
APPLICANT: Horvath, Stephen
TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
FILE REFERENCE: 689290-73
CURRENT APPLICATION NUMBER: US/09/964,824A
PRIOR FILING DATE: 2001-09-27
PRIOR APPLICATION NUMBER: US/60/236,033
PRIOR FILING DATE: 2000-09-28
PRIOR APPLICATION NUMBER: US/60/236,032
PRIOR FILING DATE: 2000-09-28
PRIOR APPLICATION NUMBER: US/60/236,028
NUMBER OF SEQ ID NOS: 583
SOFTWARE: PatentIn version 3.0
SEQ ID NO 573
LENGTH: 1804
TYPE: DNA
ORGANISM: Homo sapiens
US-09-964-824A-573

Query Match
60.0%; Score 36; DB 9; Length 1804;

Best Local Similarity 75.0%; Pred. No. 0.0036;
Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 GGTTCACGGCAGCTCCGACACGAGTCACTGTCACCCGACACCCGTCAGCTCCG 60
DB 457 GGCTCCACGGCCCCCCCCCAGCCAGGTGTCACTCGGCCCCGACACCAAGCGGCCCG 516

RESULT 12
US-10-029-517-17

Sequence 17, Application US/10029517
Publication No. US20030148969A1
GENERAL INFORMATION:
APPLICANT: Kenneth W. Dobie
APPLICANT: Susan J. Myers
TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
FILE REFERENCE: RTS-0352
CURRENT APPLICATION NUMBER: US/10/029,517
PRIOR FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 107
SEQ ID NO 17
LENGTH: 1804
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (73)...(1500)
US-10-029-517-17

Query Match
Best Local Similarity 60.0%; Score 36; DB 15; Length 1804;
Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 GGTTCACGGCAGCTCCGACACGAGTCACTGTCACCCGACACCCGTCAGCTCCG 60
DB 457 GGCTCCACGGCCCCCCCCCAGCCAGGTGTCACTCGGCCCCGACACCAAGCGGCCCG 516

RESULT 13
US-10-717-597-30

Sequence 30, Application US/10717597
Publication No. US20040110221A1
GENERAL INFORMATION:
APPLICANT: Wyeth
APPLICANT: Burczynski, Michael E.
APPLICANT: Twine, Natalie C.
APPLICANT: Dornier, Andrew J.
APPLICANT: Trepicchio, William L.
APPLICANT: Slonim, Donna K.
APPLICANT: Stover, Jennifer A.
TITLE OF INVENTION: METHODS FOR DIAGNOSING RCC AND OTHER SOLID TUMORS
FILE REFERENCE: AM010802
CURRENT APPLICATION NUMBER: US/10/717,597
PRIOR FILING DATE: 2003-11-21
PRIOR APPLICATION NUMBER: US 60/459,782
PRIOR FILING DATE: 2003-04-03
PRIOR APPLICATION NUMBER: US 60/427,982
PRIOR FILING DATE: 2002-11-21
NUMBER OF SEQ ID NOS: 4904
SOFTWARE: PatentIn version 3.2
SEQ ID NO 30
LENGTH: 1804
TYPE: DNA
ORGANISM: Homo sapiens
US-10-717-597-30

Query Match
Best Local Similarity 60.0%; Score 36; DB 17; Length 1804;
Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 GGTTCACGGCAGCTCCGACACGAGTCACTGTCACCCGACACCCGTCAGCTCCG 60
DB 457 GGCTCCACGGCCCCCCCCCAGCCAGGTGTCACTCGGCCCCGACACCAAGCGGCCCG 516

RESULT 14

US-10-775-920-84
; Sequence 84, Application US/10775920
; Publication No. US20040175744A1
; GENERAL INFORMATION:
; APPLICANT: Mergen Ltd
; TITLE OF INVENTION: POLYNUCLEOTIDE SEQUENCES AND CORRESPONDING ENCODED POLYPEPTIDES
; TITLE OF INVENTION: OF PARTICULAR SECRETED AND MEMBRANE-BOUND PROTEINS OVEREXPRESSED
; FILE REFERENCE: Mergen - 0010B
; CURRENT APPLICATION NUMBER: US/10/775,920
; CURRENT FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: US 60/447,900
; PRIOR FILING DATE: 2003-02-13
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 84
; LENGTH: 1804
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-775-920-84

Query Match 60.0%; Score 36; DB 17; Length 1804;
Best Local Similarity 75.0%; Pred. No. 0.0036;
Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 GGTTCAACGGCACTTCACGACGAGTCACTGCAACCGACCCGTCAGCTCCG 60
DB 457 GGCTCACCGCCCCCGACGCGTGTCACTCGCCCCCGACACAGCGCGCCCCG 516

RESULT 15

US-10-101-510-339
; Sequence 339, Application US/10101510
; Publication No. US20030148295A1
; GENERAL INFORMATION:
; APPLICANT: WANG, JACKSON
; APPLICANT: WANG, YIXIN
; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
; FILE REFERENCE: 15117.0012
; CURRENT APPLICATION NUMBER: US/10/101,510
; CURRENT FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/276,947
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 805
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 339
; LENGTH: 1823
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-101-510-339

Query Match 60.0%; Score 36; DB 15; Length 1823;
Best Local Similarity 75.0%; Pred. No. 0.0036;
Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 GGTTCAACGGCACTTCACGACGAGTCACTGCAACCGACCCGTCAGCTCCG 60
DB 457 GGCTCACCGCCCCCGACGCGTGTCACTCGCCCCCGACACAGCGCGCCCCG 516

Search completed: January 16, 2005, 09:30:31
Job time : 184.4 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2005, 19:50:13 : Search time 35.2 Seconds
(without alignments)
1211.572 Million cell updates/sec

Title: US-10-057-136-11

Perfect score: 60

Sequence: 1 ggtcaacggcaccctccagc.....ccgacaccgtccagctcg 60

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 824507 seqs, 35539441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Issued Patents NA:
1: /cgn2_6/prodata/1/ina/5A COMB.seq:*
2: /cgn2_6/prodata/1/ina/5B COMB.seq:*
3: /cgn2_6/prodata/1/ina/6A COMB.seq:*
4: /cgn2_6/prodata/1/ina/6B COMB.seq:*
5: /cgn2_6/prodata/1/ina/6C COMB.seq:*
6: /cgn2_6/prodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36	60.0	1804	US-10-029-517-17	Sequence 17, Appl
2	36	60.0	8186	US-10-029-517-19	Sequence 19, Appl
3	34.4	57.3	572	US-10-029-517-18	Sequence 18, Appl
4	34.4	57.3	1721	US-10-029-517-3	Sequence 3, Appl
5	33.6	56.0	518	US-10-029-517-101	Sequence 101, App
6	32.2	53.7	60	US-09-475-947A-246	Sequence 246, App
7	31.2	52.0	3343	US-10-029-517-102	Sequence 102, App
8	31.2	52.0	6192	US-08-479-537A-1	Sequence 1, Appl
9	31.2	52.0	6192	US-09-083-116-1	Sequence 1, Appl
10	31.2	52.0	6192	US-09-134-916A-1	Sequence 1, Appl
11	31.2	52.0	6449	US-08-479-537A-4	Sequence 4, Appl
12	31.2	52.0	6449	US-09-083-116-4	Sequence 4, Appl
13	31.2	52.0	6449	US-09-134-916A-4	Sequence 4, Appl
14	30.4	50.7	981	US-10-029-517-16	Sequence 16, Appl
15	28.6	47.7	519	US-09-646-028-42	Sequence 42, Appl
16	28.6	47.7	534	US-09-646-028-46	Sequence 46, Appl
17	25.4	42.3	6617	US-09-976-594-268	Sequence 268, App
18	24.6	41.0	2338	US-08-425-069-1	Sequence 1, Appl
19	24.6	41.0	2338	US-08-317-844B-1	Sequence 1, Appl
20	23.6	39.3	806	US-09-154-083-7	Sequence 7, Appl
21	23.6	39.3	2126	US-08-789-354-1	Sequence 1, Appl
22	23.6	39.3	2126	US-09-110-937-1	Sequence 1, Appl
23	23.6	39.3	2126	US-09-058-725B-1	Sequence 1, Appl
24	23.6	39.3	2126	US-09-232-857-1	Sequence 1, Appl
25	23.6	39.3	4403765	US-09-103-840A-2	Sequence 2, Appl
26	23.6	39.3	4411529	US-09-103-840A-1	Sequence 1, Appl
27	23.4	39.0	534	US-09-621-976-15293	Sequence 15293, A

c 28	23.4	39.0	1894	US-09-329-350-32	Sequence 32, Appl
c 29	23.4	39.0	1894	US-08-841-636A-32	Sequence 32, Appl
c 30	23.2	38.7	2755	US-09-907-794A-90	Sequence 90, Appl
c 31	23.2	38.7	2755	US-09-905-125A-90	Sequence 90, Appl
c 32	23.2	38.7	2755	US-09-902-775A-90	Sequence 90, Appl
c 33	23.2	38.7	2755	US-09-906-700-90	Sequence 90, Appl
c 34	23.2	38.7	2755	US-10-140-002-353	Sequence 353, App
c 35	23.2	38.7	2755	US-09-903-603A-90	Sequence 90, Appl
c 36	23	38.3	7515	US-09-328-352-1695	Sequence 1695, Ap
c 37	22.8	38.0	438	US-09-252-991A-6582	Sequence 6582, Ap
c 38	22.8	38.0	1350	US-09-252-991A-6609	Sequence 6609, Ap
c 39	22.8	38.0	1482	US-09-252-991A-6692	Sequence 6692, Ap
c 40	22.8	38.0	1644	US-09-252-991A-6563	Sequence 6563, Ap
c 41	22.8	38.0	1791	US-09-252-991A-6662	Sequence 6662, Ap
c 42	22.6	37.7	1698	US-09-484-970B-39	Sequence 887, App
c 43	22.6	37.7	3889	US-09-144-085-3	Sequence 39, Appl
c 44	22.6	37.7	33529	US-09-103-840A-2	Sequence 3, Appl
c 45	22.6	37.7	4403765	US-09-103-840A-2	Sequence 2, Appl

ALIGNMENTS

```

RESULT 1
US-10-029-517-17
Sequence 17, Application US/10029517
Patent No. 6716627
GENERAL INFORMATION:
APPLICANT: Kenneth W. Dobie
APPLICANT: Susan J. Myers
TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
FILE REFERENCE: RTS-0352
CURRENT APPLICATION NUMBER: US/10/029,517
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 107
SEQ ID NO 17
LENGTH: 1804
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (73)...(1500)
US-10-029-517-17

Query Match      60.0%; Score 36; DB 4; Length 1804;
Best Local Similarity 75.0%; Pred. No. 0.0021;
Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY      1  GGTCAACGGCACCCTCCAGCAGAGTCAGCTGCACCGCAGACCCGTCAGCTCGG 60
Db      457 GGTCTCACCGCCCCCCCCCAGCAGCTGTCAGCTCGCCCCGACACAGCGCGCCCG 516

RESULT 2
US-10-029-517-19
Sequence 19, Application US/10029517
Patent No. 6716627
GENERAL INFORMATION:
APPLICANT: Kenneth W. Dobie
APPLICANT: Susan J. Myers
TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
FILE REFERENCE: RTS-0352
CURRENT APPLICATION NUMBER: US/10/029,517
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 107
SEQ ID NO 19
LENGTH: 8186
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: unsure
LOCATION: 6899
OTHER INFORMATION: unknown

```

```

; NAME/KEY: unsure
; LOCATION: 7155
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 7184
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 7957
; OTHER INFORMATION: unknown
; NAME/KEY: intron
; LOCATION: (2997)...(3498)
; OTHER INFORMATION: intron 1
; NAME/KEY: intron:exon junction
; LOCATION: (3498)...(3499)
; OTHER INFORMATION: intron 1:exon 2
; NAME/KEY: exon
; LOCATION: (3508)...(3599)
; OTHER INFORMATION: exon 2d
; NAME/KEY: exon:intron junction
; LOCATION: (3982)...(3983)
; OTHER INFORMATION: exon 2a:intron 2a
; NAME/KEY: intron:exon junction
; LOCATION: (4205)...(4206)
; OTHER INFORMATION: intron 2c:exon 3c
; NAME/KEY: intron:exon junction
; LOCATION: (4259)...(4260)
; OTHER INFORMATION: intron 2d:exon 3d
; NAME/KEY: exon
; LOCATION: (4260)...(4328)
; OTHER INFORMATION: exon 3d
; NAME/KEY: intron:exon junction
; LOCATION: (4632)...(4633)
; OTHER INFORMATION: intron 3:exon 4
; NAME/KEY: exon
; LOCATION: (4914)...(5035)
; OTHER INFORMATION: exon 5
; NAME/KEY: intron
; LOCATION: (5266)...(6293)
; OTHER INFORMATION: intron 6
US-10-029-517-19

```

Query Match
Best Local Similarity 60.0%; Score 36; DB 4; Length 8186;
Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

```

Qy 1 GGTTCACGGCAGCCTCCAGACACGAGTCACTGTGACACCCGAGTCCAGCTCG 60
Db 3825 GGCTCCACCGCCGCCCGCCAGCCAGGTGTCACTCGGCCGAGACACGAGCCGCCCG 3864

```

```

RESULT 3
US-10-029-517-18
; Sequence 18, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 18
; LENGTH: 572
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (67)...(572)
US-10-029-517-18

```

Query Match
Best Local Similarity 57.3%; Score 34.4; DB 4; Length 572;
Matches 73.3%; Pred. No. 0.0059;

```

Matches 44; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
Qy 1 GGTTCACGGCAGCCTCCAGACACGAGTCACTGTGACACCCGAGTCCAGCTCG 60
Db 478 GGCTCCACCGCCGCCCGCCAGCCAGGTGTCACTCGGCCGAGACACGAGCCGCCCG 537

```

```

RESULT 4
US-10-029-517-3
; Sequence 3, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 3
; LENGTH: 1721
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (58)...(1605)
US-10-029-517-3

```

Query Match
Best Local Similarity 57.3%; Score 34.4; DB 4; Length 1721;
Matches 44; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

```

Qy 1 GGTTCACGGCAGCCTCCAGACACGAGTCACTGTGACACCCGAGTCCAGCTCG 60
Db 442 GGCTCCACCGCCGCCCGCCAGCCAGGTGTCACTCGGCCGAGACACGAGCCGCCCG 501

```

```

RESULT 5
US-10-029-517-101
; Sequence 101, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 101
; LENGTH: 518
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (101)...(518)
US-10-029-517-101

```

Query Match
Best Local Similarity 56.0%; Score 33.6; DB 4; Length 518;
Matches 42; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

```

Qy 1 GGTTCACGGCAGCCTCCAGACACGAGTCACTGTGACACCCGAGTCCAGCTCG 56
Db 462 GGCTCCACCGCCGCCCGCCAGCCAGGTGTCACTCGGCCGAGACACGAGCCGCCCG 517

```

```

RESULT 6
US-09-475-947A-246
; Sequence 246, Application US/09475947A
; Patent No. 6472154
; GENERAL INFORMATION:
; APPLICANT: Garner, Harold R.
; APPLICANT: Wren, Jonathan D.
; APPLICANT: Minna, John D.

```



```
/ TITLE OF INVENTION: Polymorphic Repeats in Human Genes
/ FILE REFERENCE: UTS0667
/ CURRENT APPLICATION NUMBER: US/09/475,947A
/ CURRENT FILING DATE: 1999-12-31
/ NUMBER OF SEQ ID NOS: 346
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 246
/ LENGTH: 60
/ TYPE: DNA
/ ORGANISM: human
US-09-475-947A-246

Query Match
Best Local Similarity 75.5%; Score 32.2; DB 4; Length 60;
Matches 40; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 GGTTCACGGCAGCTTCAGACAGGAGTGCAGCTGCAACCCGACACCGCTCC 53
DB 7 GGCTCCACCGCCCCCGCCAGCCGCGTGTCACTCGCCCGGACACGAGCC 59

RESULT 7
US-10-029-517-102
/ Sequence 102, Application US/10029517
/ Patent No. 6716627
/ GENERAL INFORMATION:
/ APPLICANT: Kenneth W. Dobie
/ APPLICANT: Susan J. Myers
/ TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
/ FILE REFERENCE: RTS-0352
/ CURRENT APPLICATION NUMBER: US/10/029,517
/ CURRENT FILING DATE: 2001-12-20
/ NUMBER OF SEQ ID NOS: 107
/ SEQ ID NO 102
/ LENGTH: 3343
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-029-517-102

Query Match
Best Local Similarity 70.0%; Score 31.2; DB 4; Length 3343;
Matches 42; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1 GGTTCACGGCAGCTTCAGACAGGAGTGCAGCTGCAACCCGACACCGCTCCG 60
DB 1608 GGCTCCACCGCCCCCGCCAGCCGCGTGTCACTCGCCCGGACAAAGCCAGCCCG 1667

RESULT 8
US-08-479-537A-1
/ Sequence 1, Application US/08479537A
/ Patent No. 5861381
/ GENERAL INFORMATION:
/ APPLICANT: CHAMBERON, Pierre
/ APPLICANT: KIENEY, Marie-Paule
/ APPLICANT: LATHE, Richard
/ APPLICANT: HAREDEVANT, Maira
/ TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
/ TITILE OF INVENTION: TREATMENT OR PREVENTION OF A MALIGNANT TUMOR
/ NUMBER OF SEQUENCES: 5
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
/ STREET: P.O. Box 1404
/ CITY: Alexandria
/ STATE: Virginia
/ COUNTRY: United States
/ ZIP: 22313-1404
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
```

```
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/479,537A
/ FILING DATE: 07-JUN-1995
/ CLASSIFICATION: 514
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: FR 90/13101
/ FILING DATE: 23-OCT-1990
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: WO PCT/FR91/00835
/ FILING DATE: 23-OCT-1991
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/039,320
/ FILING DATE: 04-APR-1993
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/403,576
/ FILING DATE: 14-MAR-1995
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Teskin, Robin L.
/ REGISTRATION NUMBER: 35,030
/ REFERENCE/DOCKET NUMBER: 017753-025
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (703) 836-6620
/ TELEFAX: (703) 836-2021
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 6192 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
/ FEATURE:
/ NAME/KEY: sig_peptide
/ LOCATION: 58..120
/ FEATURE:
/ NAME/KEY: repeat_region
/ LOCATION: 439..5239
/ OTHER INFORMATION: /note="The nucleotides spanning
/ OTHER INFORMATION: 439-5239 constitute a repeated region wherein the repeat is 6
/ OTHER INFORMATION: nucleotides and encodes 20 amino acids, 17 of which are fixed
/ OTHER INFORMATION: The number of such repeats varies from 1 to 80."
/ FEATURE:
/ NAME/KEY: mat_peptide
/ LOCATION: 121..6166
/ FEATURE:
/ NAME/KEY: repeat_region
/ LOCATION: 457
/ OTHER INFORMATION: /note="Nucleotide 457 is X1 = NNN
/ OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
/ OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
/ FEATURE:
/ NAME/KEY: repeat_region
/ LOCATION: 487
/ OTHER INFORMATION: /note="Nucleotide 487 is Y = NNN
/ OTHER INFORMATION: which is the codon for Thr or Asn wherein Thr = ACT, ACC, ACA
/ OTHER INFORMATION: or ACG; and Asn = AAT or AAC."
/ FEATURE:
/ NAME/KEY: repeat_region
/ LOCATION: 496
/ OTHER INFORMATION: /note="Nucleotide 496 is X2 = NNN
/ OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
/ OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
US-08-479-537A-1

Query Match
Best Local Similarity 70.0%; Score 31.2; DB 2; Length 6192;
Matches 42; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1 GGTTCACGGCAGCTTCAGACAGGAGTGCAGCTGCAACCCGACACCGCTCCG 60
DB 382 GGCTCCACCGCCCCCGCCAGCCGCGTGTCACTCGCCCGGACAAAGCCAGCCCG 441

RESULT 9
```

US-09-083-116-1
; Sequence 1, Application US/09083116
; Patent No. 6283795
; GENERAL INFORMATION:
; APPLICANT: CHAMON, Pierre
; APPLICANT: KIENY, Marie-Paule
; APPLICANT: LATHE, Richard
; APPLICANT: HAREUVENI, Mara
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
; TREATMENT OR PREVENTION OF A MALIGNANT TUMOR
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/083,116
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/479,537
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/FR91/00835
; FILING DATE: 23-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/039,320
; FILING DATE: 04-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/403,576
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Teskin, Robin L.
; REGISTRATION NUMBER: 35,030
; REFERENCE/DOCKET NUMBER: 017753-025
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6192 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 58..120
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 439..5239
; OTHER INFORMATION: /note="The nucleotides spanning
; 439-5239 constitute a repeated region wherein the repeat is 6
; OTHER INFORMATION: nucleotides and encodes 20 amino acids, 17 of which are fixed
; OTHER INFORMATION: The number of such repeats varies from 1 to 80."
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 121..6166
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 457
; OTHER INFORMATION: /note="Nucleotide 457 is X1 = NNN
; OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
; OTHER INFORMATION: or CCG, and Ala = GCT, GCC, GCA, or GCG."
; FEATURE:
; OR CCG, and Ala = GCT, GCC, GCA, or GCG."

NAME/KEY: repeat_region
LOCATION: 487
OTHER INFORMATION: /note="Nucleotide 487 is Y = NNN
OTHER INFORMATION: which is the codon for Thr or Asn wherein Thr = ACT, ACC, ACA
OTHER INFORMATION: or ACG, and Asn = AAT or AAC."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 496
OTHER INFORMATION: /note="Nucleotide 496 is X2 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
OTHER INFORMATION: or CCG, and Ala = GCT, GCC, GCA, or GCG."
US-09-083-116-1
Query Match 52.0%; Score 31.2; DB 3; Length 6192;
Best Local Similarity 70.0%; Pred. No. 0.11;
Matches 42; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
QY 1 GGTCACGCGACCCGACGAGTCACGTCGACCCGACACCCGTCGAGTCGCG 60
DB 382 GGCTCACACCCCGCCGACGATGTCACCTGACCCCGGACAAACAGCCGCCCC 441
RESULT 10
US-09-134-916A-1
; Sequence 1, Application US/09134916A
; Patent No. 6328956
; GENERAL INFORMATION:
; APPLICANT: CHAMON, Pierre
; APPLICANT: KIENY, Marie-Paule
; APPLICANT: LATHE, Richard
; APPLICANT: HAREUVENI, Mara
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
; TREATMENT OR PREVENTION OF A MALIGNANT TUMOR
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/134,916A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/479,537
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: FR 90/13101
; FILING DATE: 23-OCT-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/FR91/00835
; FILING DATE: 23-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/039,320
; FILING DATE: 04-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/403,576
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Teskin, Robin L.
; REGISTRATION NUMBER: 35,030
; REFERENCE/DOCKET NUMBER: 017753-025
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:

```

US-09-083-116-4
Sequence 4, Application US/09083116

RESULT 12

FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 90/13101
FILING DATE: 23-OCT-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/FR91/00835
FILING DATE: 23-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/039,320
FILING DATE: 04-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/403,576
FILING DATE: 14-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Teakin, Robin L.
REGISTRATION NUMBER: 35,030
REFERENCE/DOCKET NUMBER: 017753-025
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 6449 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 56..120
FEATURE:
NAME/KEY: repeat_region
LOCATION: 439..5239
OTHER INFORMATION: /note= "The nucleotides spanning 439-5239 constitute a repeated region wherein the repeat is { OTHER INFORMATION: nucleotides and encodes 20 amino acids, 17 of which are fixed OTHER INFORMATION: The number of such repeats varies from 1 to 80."
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 121..5661
FEATURE:
NAME/KEY: repeat_region
LOCATION: 457
OTHER INFORMATION: /note= "Nucleotide 457 is X1 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCf
OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 487
OTHER INFORMATION: /note= "Nucleotide 487 is Y = NNN
OTHER INFORMATION: which is the codon for Thr or Asn wherein Thr = ACT, ACC, ACf
OTHER INFORMATION: or ACG; and Asn = AAT or AAC."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 496
OTHER INFORMATION: /note= "Nucleotide 496 is X2 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCf
OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
US-08-479-537A-4

Query Match 52.0%; Score 31.2; DB 2; Length 6449;
Best Local Similarity 70.0%; Pred. No. 0.11;
Matches 42; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

1 GGTTCAAGCCGACACTTCAGACACGAGTCAAGTCTGCACCCGCCGCTCAGGTCCG 60
Db GGCTCCACCACCCCCTGGCAGCCAGATGTCACCTCAGCCCCGAGCAAAAGCACGCCCG 441

```

Patent No. 6203795
GENERAL INFORMATION:
APPLICANT: CHAMON, Pierre
APPLICANT: KIENY, Marie-Paule
APPLICANT: LATHE, Richard
APPLICANT: HAREVENT, Mara
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
TREATMENT OR PREVENTION OF A MALIGNANT TUMOR
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/083,116
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/479,537
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/FR91/00835
FILING DATE: 23-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/039,320
FILING DATE: 04-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/403,576
FILING DATE: 14-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Teskin, Robin L.
REGISTRATION NUMBER: 35,030
REFERENCE/DOCKET NUMBER: 017753-025
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 6449 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: s19_peptide
LOCATION: 58..120
FEATURE:
NAME/KEY: repeat_region
LOCATION: 439..5238
OTHER INFORMATION: /note= "The nucleotides spanning
439-5239 constitute a repeated region wherein the repeat is 6
nucleotides and encodes 20 amino acids, 17 of which are fixed
OTHER INFORMATION: The number of such repeats varies from 1 to 80."
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 121..5661
FEATURE:
NAME/KEY: repeat_region
LOCATION: 457
OTHER INFORMATION: /note= "Nucleotide 457 is XI = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
FEATURE: or CCG; and Ala = GCT, GCC, GCA, or GCG."
NAME/KEY: repeat_region
LOCATION: 487

OTHER INFORMATION: /note= "Nucleotide 487 is Y = NNN
OTHER INFORMATION: which is the codon for Thr or Asn wherein Thr = ACT, ACC, ACA
OTHER INFORMATION: or ACG; and Asn = AAT or AAC."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 456
OTHER INFORMATION: /note= "Nucleotide 496 is X2 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
US-09-083-116-4
Query Match 52.0%; Score 31.2; DB 3; Length 6449;
Best Local Similarity 70.0%; Pred. No. 0.11;
Matches 42; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
QY 1 GGTTCACGCGACCTCCAGCAGCAGAGTCAGTCGACCCGACCGCTCCAGCTCCG 60
DB 382 GGCTCCACACCCCGCCGACGATGTCTACCTCAGCCCGGACACAGACGACCGCCG 441
RESULT 13
US-09-134-916A-4
Sequence 4, Application US/09134916A
Patent No. 6328956
GENERAL INFORMATION:
APPLICANT: CHAMON, Pierre
APPLICANT: KIENY, Marie-Paule
APPLICANT: LATHE, Richard
APPLICANT: HAREVENT, Mara
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
TREATMENT OR PREVENTION OF A MALIGNANT TUMOR
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/134,916A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/479,537
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: FR 90/13101
FILING DATE: 23-OCT-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/FR91/00835
FILING DATE: 23-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/039,320
FILING DATE: 04-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/403,576
FILING DATE: 14-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Teskin, Robin L.
REGISTRATION NUMBER: 35,030
REFERENCE/DOCKET NUMBER: 017753-025
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 6449 base pairs
TYPE: nucleic acid

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2005, 19:43:55 ; Search time 1550.6 Seconds
(without alignments)
1410.024 Million cell updates/sec

Title: US-10-057-136-11

Perfect score: 60
Sequence: 1 ggctcaacgcacacccacgc.....ccgacacccgcagctccg 60

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hnc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_g881:*
9: gb_g882:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	37.6	62.7	604	4	BM791359 K-EST0071
2	37.6	62.7	877	5	BUS42454 AGENCOURT
3	37.6	62.7	959	6	CA489836 AGENCOURT
4	37.6	62.7	1113	5	BUI48487 AGENCOURT
5	37.6	62.7	1234	5	BQ936898 AGENCOURT
6	37.6	62.7	1262	5	BQ935496 AGENCOURT
7	37.6	62.7	1343	5	BQ920055 AGENCOURT
8	37.6	62.7	1531	5	BUS43309 AGENCOURT
9	37.6	60.0	330	1	AI925867 w02004.x
10	36	60.0	1130	5	BUS42996 AGENCOURT
11	36	60.0	1268	5	BQ943554 AGENCOURT
12	36	60.0	1349	5	BUI52566 AGENCOURT
13	36	60.0	1420	5	BUS42790 AGENCOURT
14	31.2	52.0	472	4	BM759495 K-EST0039
15	30.4	50.7	981	4	BG774910 602649832
16	30.4	50.7	1334	5	BQ943809 AGENCOURT
17	30.4	50.7	1536	5	BQ923149 AGENCOURT
18	29.6	49.3	1678	4	BG75565 602650481
19	29.4	49.0	166	7	T27692 EST12384 HU
20	29.4	49.0	475	6	CB120860 K-EST0168
21	29.4	49.0	669	6	CB122585 K-EST0170
22	29	48.3	619	4	BI260921 602970962
23	28.8	48.0	932	4	BG444776 GA_Ea002
24	28.6	47.7	1279	9	AG097186 Pan trogl

C 25	27.6	46.0	356	8	AQ906176 GSSTC0667
C 26	27.6	46.0	395	4	BM098904 EBD105 SQ
C 27	27.4	45.7	540	1	AJ470277 AJ470277
C 28	27.2	45.3	528	4	BM130288 pb29D06.Y
C 29	27.2	45.3	538	4	BI744470 pD09C12.Y
C 30	27.2	45.3	547	8	AQ841698 t137709b
C 31	27.2	45.3	739	9	AG517386 nus nubcu
C 32	27.2	45.3	762	5	BX875037 BX875037
C 33	27.2	45.3	768	9	CC901252 ZMMBB023
C 34	27.2	45.3	787	9	CC877595 ZMMBB021
C 35	27.2	45.3	893	4	BM045068 603622933
C 36	27.2	45.3	1231	5	BQ890022 AGENCOURT
C 37	27	45.0	240	1	AJ460645 AJ460645
C 38	27	45.0	262	1	AV919837 AV919837
C 39	27	45.0	300	1	AJ460651 AJ460651
C 40	27	45.0	313	4	BI776854 EBD103 SQ
C 41	27	45.0	314	8	AZ227363 RPT1-23-7
C 42	27	45.0	315	4	BM097405 EBem07 SQ
C 43	27	45.0	317	6	CB862996 HH01P0Bu
C 44	27	45.0	323	6	CA028771 HZ63D10X
C 45	27	45.0	324	6	CB876471 HX11G01W

ALIGNMENTS

RESULT 1
BM791359
LOCUS 604 bp mRNA linear EST 05-MAR-2002
DEFINITION K-EST0071342 S21SNUS20 Homo sapiens cDNA clone S21SNUS20-14-A06 5',
mRNA sequence.
ACCESSION BM791359
VERSION BM791359.1 GI:19139591
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 604)
Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
21C Frontier Korean EST Project 2001
Unpublished (2002)
JOURNAL
CONTACT: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 14 row: A column: 06
High quality sequence stop: 604.
Location/Qualifiers
1. 604
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S21SNUS20-14-A06"
/sex="F"
/tissue_type="Stomach"
/cell_type="Floating aggregates"
/cell_line="SN0-520"
/lab_host="Top10F"
/clone_lib="S21SNUS20"
/note="Organ: Stomach; Vector: pTZ19RP; Site 1: EcorI;
Site 2: NotI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then deapped
with tobacco acid pyrophosphatase (TAP). The deapped
Intact mRNA was ligated with DNA-RNA linker including Ecor
I site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized from oligo dt-selected mRNA by
priming with dt-tailed vector. The dt-tailed vector was

adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F⁺ by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

ORIGIN

Query Match 62.7%; Score 37.6; DB 4; Length 604;
Best Local Similarity 76.7%; Pred. No. 0.12;
Matches 46; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 1 GGTTCACGGCAGCTCCAGACACGAGTCACTGACACCGGACCGCTCCAGCTCCG 60
28 GGCTCCACCGCAGCCCGCCAGCCGAGTGTACCTGGCCCGGACACGAGCGGCGCCG 87

RESULT 2
BU542454 877 bp mRNA linear EST 13-SEP-2002
LOCUS AGENCOURT 10322173 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6574322

DEFINITION 5', mRNA sequence.
ACCESSION BU542454
VERSION BU542454.1 GI:22852937

KEYWORDS EST.
SOURCE Homo sapiens (human)

ORGANISM

REFERENCE 1 (bases 1 to 877)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DTP

cDNA Library Preparation: Rubin Laboratory
DNA Sequencing Arrayed by: The I.M.A.G.E. Consortium (LNL)
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LNCM2769 row: h column: 02
High quality sequence stop: 760.
Location/Qualifiers

FEATURES

Source

1. 877
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6574322"
/tissue_type="carcinoma, cell line"
/lab_host="NIH_MGC_40"
/note="Organ: prostate; Vector: pORF7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(6). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 62.7%; Score 37.6; DB 5; Length 877;
Best Local Similarity 76.7%; Pred. No. 0.12;
Matches 46; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 1 GGTTCACGGCAGCTCCAGACACGAGTCACTGACACCGGACCGCTCCAGCTCCG 60
215 GGCTCCACCGCAGCCCGCCAGCCGAGTGTACCTGGCCCGGACACGAGCGGCGCCG 274

RESULT 3
CA489836 959 bp mRNA linear EST 14-NOV-2002
LOCUS AGENCOURT 10810668 MAPCL Homo sapiens cDNA clone IMAGE:6722324 5',
DEFINITION mRNA sequence.
ACCESSION CA489836
VERSION CA489836.1 GI:24952627

KEYWORDS EST.
SOURCE Homo sapiens (human)

ORGANISM

REFERENCE 1 (bases 1 to 959)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Kristi A. Egland, Ira Pastan

cDNA Library Preparation: Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM14284 row: n column: 20
High quality sequence start: 31
High quality sequence stop: 446.
Location/Qualifiers

FEATURES

Source

1. 959
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6722324"
/cell_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231, hTERT-HME1, LNCaP"
/lab_host="EMD10B"
/note="Vector: pCMV-SPORT6; Site 1: EcoRV; Site 2: Not I; Subcloned with brain, liver, lung, kidney and muscle. Directionally cloned. Priming method: oligo-dT. Average insert size: 1800 bp. Library amplification: 26,000 fold. Kristi A. Egland, James J. Vincent, Robert Strausberg, Bungkok Lee & Ira Pastan: Discovery of new breast cancer genes encoding membrane and secreted proteins. Manuscript submitted."

ORIGIN

Query Match 62.7%; Score 37.6; DB 6; Length 959;
Best Local Similarity 76.7%; Pred. No. 0.12;
Matches 46; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 1 GGTTCACGGCAGCTCCAGACACGAGTCACTGACACCGGACCGCTCCAGCTCCG 60
469 GGCTCCACCGCAGCCCGCCAGCCGAGTGTACCTGGCCCGGACACGAGCGGCGCCG 528

RESULT 4

BU148487 1113 bp mRNA linear EST 03-SEP-2002
LOCUS AGENCOURT 8670479 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6380642
DEFINITION 5', mRNA sequence.
ACCESSION BU148487
VERSION BU148487.1 GI:22662019

KEYWORDS EST.
SOURCE Homo sapiens (human)

ORGANISM

REFERENCE 1 (bases 1 to 1113)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL
COMMENT

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLCM2569 row: j column: 03
High quality sequence stop: 235.

FEATURES

Location/Qualifiers
1. 1113

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6380642"
/issue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 62.7%; Score 37.6; DB 5; Length 1113;
Best Local Similarity 76.7%; Pred. No. 0.12; Indels 0; Gaps 0;
Matches 46; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 1 GGTTCAACGGACCTTCACGACGAGTCAGTCCGACACCCGTCAGCTCCG 60
DB 30 GGCTCACGGCCCCCGACGCGGTGCTGCTGCCCCGACACGAGCGGCCCG 89

RESULT 5
BO936898 1234 bp mRNA linear EST 21-AUG-2002
LOCUS AGENCOURT_8919110 NIH_MGC_40 Homo sapiens CDNA clone IMAGE:6383855
DEFINITION 5', mRNA sequence.
ACCESSION BO936898
VERSION BO936898.1 GI:22352281
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE NIH-MGC http://mgi.nci.nih.gov/1 (bases 1 to 1234)
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLCM2577 row: o column: 24
High quality sequence stop: 245.

FEATURES

Location/Qualifiers
1. 1234

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6383855"

ORIGIN

/issue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

Query Match 62.7%; Score 37.6; DB 5; Length 1234;
Best Local Similarity 76.7%; Pred. No. 0.12; Indels 0; Gaps 0;
Matches 46; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 1 GGTTCAACGGACCTTCACGACGAGTCAGTCCGACACCCGTCAGCTCCG 60
DB 30 GGCTCACGGCCCCCGACGCGGTGCTGCTGCCCCGACACGAGCGGCCCG 89

RESULT 6
BO935496 1262 bp mRNA linear EST 21-AUG-2002
LOCUS AGENCOURT_876475 NIH_MGC_40 Homo sapiens CDNA clone IMAGE:6382088
DEFINITION 5', mRNA sequence.
ACCESSION BO935496
VERSION BO935496.1 GI:22350879
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE NIH-MGC http://mgi.nci.nih.gov/1 (bases 1 to 1262)
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLCM2573 row: f column: 09
High quality sequence start: 46
High quality sequence stop: 157.

FEATURES

Location/Qualifiers
1. 1262

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6382088"
/issue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 62.7%; Score 37.6; DB 5; Length 1262;
Best Local Similarity 76.7%; Pred. No. 0.12; Indels 0; Gaps 0;
Matches 46; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2455879"
/issue_type="adenocarcinoma"
/lab_host="DH10B"
/note="Organ: pancreas; Vector: pCW-Sport6; Site 1: SalI; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.72 Kb. Life Technologies catalog #: 11548-013"

ORIGIN

Query Match 60.0%; Score 36; DB 1; Length 330;
Best Local Similarity 75.0%; Pred. No. 0.36;
Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Oy 1 GGTTCACGGCAGCTCCAGACAGGAGTCAGCTGACACCCGACCCGTCAGTCCG 60
|||
Db 316 GGCTCACGCGCCCCCAGCCAGCGGTGTCACCTCGCCCCGAGACACGAGCGCCCG 257

RESULT 10 1130 bp mRNA linear EST 13-SEP-2002
LOCUS BU542996
DEFINITION AGENCOURT_10338707 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6574979
5', mRNA sequence.
ACCESSION BU542996
VERSION BU542996.1 GI:22853479
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 1130)
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: L1CM2771 row: c column: 11
High quality sequence start: 27
High quality sequence stop: 246.
Location/Qualifiers

FEATURES

1..1130
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6574979"
/issue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_1ib="NIH MGC 40"
/note="Organ: prostate; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dt priming.
Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN

Query Match 60.0%; Score 36; DB 5; Length 1130;
Best Local Similarity 75.0%; Pred. No. 0.39;
Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Oy 1 GGTTCACGGCAGCTCCAGACAGGAGTCAGTGTGACACCCGACCCGTCAGTCCG 60
|||
Db 69 GGCTCACGCGCCCCCAGCCAGCGGTGTCACCTCGCCCCGAGACACGAGCGCCCG 128

RESULT 11 1268 bp mRNA linear EST 21-AUG-2002
LOCUS BO943554
DEFINITION AGENCOURT_8777815 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6384308
5', mRNA sequence.
ACCESSION BO943554
VERSION BO943554.1 GI:22359032
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 1268)
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: L1CM2579 row: b column: 21
High quality sequence stop: 177.
Location/Qualifiers

FEATURES

1..1268
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6384308"
/issue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_1ib="NIH MGC 40"
/note="Organ: prostate; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dt priming.
Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN

Query Match 60.0%; Score 36; DB 5; Length 1268;
Best Local Similarity 75.0%; Pred. No. 0.39;
Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Oy 1 GGTTCACGGCAGCTCCAGACAGGAGTCAGTGTGACACCCGACCCGTCAGTCCG 60
|||
Db 90 GGCTCACGCGCCCCCAGCCAGCGGTGTCACCTCGCCCCGAGACACGAGCGCCCG 149

RESULT 12 1349 bp mRNA linear EST 03-SEP-2002
LOCUS BU152566
DEFINITION AGENCOURT_8670683 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6380559
5', mRNA sequence.
ACCESSION BU152566
VERSION BU152566.1 GI:22666098
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 1349)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-r@mail.nih.gov
 Tissue Procurement: DCTD/DRP

CDNA Library Preparation: Rubin Laboratory
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: L1CM2569 row: f column: 16
 High quality sequence stop: 291.
 Location/Qualifiers

FEATURES

source

1..1349
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6380559"
 /tissue_type="carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_40"
 /note="Organ: prostate; Vector: pOTB7; Site 1: XhoI;
 Site 2: EcoRI; CDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCACGAG(G). Library constructed by
 Ling Hong in the laboratory of Gerald M. Rubin (University
 of California, Berkeley) using ZAP-CDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library."

ORIGIN

Query Match 60.0%; Score 36; DB 5; Length 1349;
 Best Local Similarity 75.0%; Pred. No. 0.39;
 Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 GGTTCACGGCACTCTCAGACACGAGTCACTGTGCACCCGACCCGTCGAGTCCG 60
 19 GGCTCACCGCCGCCCGCCAGCCAGGTGTCACTCGGCCCGGACACGAGCCGCCCG 78

RESULT 13 1420 bp mRNA linear EST 13-SHP-2002
 LOCUS B542790
 DEFINITION AGENCOURT_10334841 NIH_MGC_40 Homo sapiens CDNA clone IMAGE:6574725
 5', mRNA sequence.
 ACCESSION B542790
 VERSION B542790.1 GI:22853273
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 1 (bases 1 to 1420)
 NIH-MGC <http://mgc.nci.nih.gov/>
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-r@mail.nih.gov
 Tissue Procurement: DCTD/DRP

CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: L1CM2770 row: h column: 21
 High quality sequence stop: 288.
 Location/Qualifiers

FEATURES

source

1..1420
 /organism="Homo sapiens"
 /mol_type="mRNA"

ORIGIN

Query Match 60.0%; Score 36; DB 5; Length 1420;
 Best Local Similarity 75.0%; Pred. No. 0.39;
 Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 GGTTCACGGCACTCTCAGACACGAGTCACTGTGCACCCGACCCGTCGAGTCCG 60
 19 GGCTCACCGCCGCCCGCCAGCCAGGTGTCACTCGGCCCGGACACGAGCCGCCCG 78

RESULT 14 472 bp mRNA linear EST 04-MAR-2002
 LOCUS BMT59495
 DEFINITION K-EST0039670 S9SNU601 Homo sapiens CDNA clone S9SNU601-2-F03 5',
 mRNA sequence.
 ACCESSION BMT59495
 VERSION BMT59495.1 GI:19089110
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 1 (bases 1 to 472)
 Oh, K.Y., Hahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R.,
 Kim, Y.S., Cheong, J.B., Sohn, H.Y., Kim, V.M., Park, H.S., Kim, S. and
 Kim, Y.S.
 21C Frontier Korean EST Project 2001
 Unpublished (2002)
 Contact: Kim YS
 Genome Research Center
 Korea Research Institute of Bioscience & Biotechnology
 52 Eosun-dong Yuseong-gu, Daejeon 305-333, South Korea
 Tel.: +82-42-860-4470
 Fax: +82-42-860-4409
 Email: yongsung@mail.krrib.re.kr
 Plate: 2 row: F column: 03
 High quality sequence stop: 472.
 Location/Qualifiers

FEATURES

source

1..472
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="S9SNU601-2-F03"
 /sex="W"
 /tissue_type="testes"
 /cell_type="epithelial"
 /cell_line="SNU-601"
 /lab_host="Top10P"
 /clone_lib="S9SNU601"
 /note="Organ: Stomach; Vector: pMR18-FU3; Site 1: XhoI;
 Site 2: XhoI; The poly (A)+ RNA was dephosphorylated with
 bacterial alkaline phosphatase (BAP) and then deacapped
 with tabacco acid pyrophosphatase (TAP). The deacapped
 intact mRNA was ligated with DNA-RNA linker including SfiI
 site by treatment of T4 RNA ligase and the first strand
 cDNA was synthesized with Superscript II using SfiI
 oligo-dT primer. After first strand synthesis RNA was
 degraded by NaOH treatment and cDNA was amplified by PCR
 reaction. The PCR products were digested with SfiI and

ORIGIN

cloned into DraIII-digested pME18S-Fli3 vector. The obtained cDNA vectors were used for transformation of competent cells *E. coli* Top10F' by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

Query Match 52.0%; Score 31.2; DB 4; Length 472;
Best Local Similarity 70.0%; Pred. No. 13;
Matches 42; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1 GGTTCAACGGCAGCTCCAGACAGGAGTCAGCTGACACCCGACCGCTCCGCG 60
DB 391 GGCTCCACGACCCCGCAGCCGACGATGTCACCTCAGCCCGGACAAAGCCGCGG 450

RESULT 15

BG774910 981 bp mRNA linear EST 15-MAY-2001
LOCUS 602649832F1 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:4761054.5',
DEFINITION mRNA sequence.

ACCESSION BG774910 GI:14045227
VERSION BG774910
KEYWORDS EST.

SOURCE Homo sapiens (human)
ORGANISM

REFERENCE Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE NIH-MGC http://mgs.nci.nih.gov/.
JOURNAL 1 (bases 1 to 981)
COMMENT Unpublished (1999)
National Institutes of Health, Mammalian Gene Collection (MGC)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LNCM1612 row: 0 column: 07
High quality sequence stop: 874.

FEATURES

source

Location/Qualifiers
1..981
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4761054"
/feature_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCAAGG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC library."

ORIGIN

Query Match 50.7%; Score 30.4; DB 4; Length 981;
Best Local Similarity 71.4%; Pred. No. 23;
Matches 40; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1 GGTTCAACGGCAGCTCCAGACAGGAGTCAGCTGACACCCGACCGCTCCGCG 56
DB 21 GGCTCCACGACCCCGCAGCCGACGATGTCACCTCAGCCCGGACAAAGCCGCG 76

Search completed: January 16, 2005, 02:55:40
Job time : 1552.6 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2005, 16:00:11 ; Search time 170.5 Seconds
(without alignments)
1847.304 Million cell updates/sec

Title: US-10-057-136-11

Perfect score: 60

Sequence: 1 ggtcacacgcacccacgcacg.....ccgacacccgcgcagctccg 60

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : N_Geneseq_23Sep04:*
1: genebegn1980s:*
2: genebegn1990s:*
3: genebegn2000s:*
4: genebegn2001as:*
5: genebegn2001bs:*
6: genebegn2002as:*
7: genebegn2002bs:*
8: genebegn2003as:*
9: genebegn2003bs:*
10: genebegn2003cs:*
11: genebegn2003ds:*
12: genebegn2004s:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	60	2	AAV48324 Nucleotide
2	40.8	68.0	156	10	ADK68635 HSP65-MUC
3	39.2	65.3	1800	10	ADK68629 HSP65-MUC
4	38.2	63.7	60	2	AAV48318 Nucleotide
5	36.6	61.0	60	2	AAV48319 Nucleotide
6	36	60.0	60	2	AAV48316 Nucleotide
7	36	60.0	78	2	AAV48326 Nucleotide
8	36	60.0	309	1	AAV90579 PDF9.3 CD
9	36	60.0	1194	12	AD157712 Human bre
10	36	60.0	1378	12	AD157693 Human bre
11	36	60.0	1424	12	AD023180 Antisense
12	36	60.0	1428	6	ABL60159 Human MUC
13	36	60.0	1428	12	AD023125 Human MUC
14	36	60.0	1457	12	ADP32627 Plasmid J
15	36	60.0	1527	2	AAV48329 Human MUC
16	36	60.0	1614	12	ADK70370 Respirato
17	36	60.0	1630	12	AD157708 Human bre
18	36	60.0	1634	12	AD157689 Human bre
19	36	60.0	1712	12	AD157686 Human bre
20	36	60.0	1738	12	AD157669 Human bre
21	36	60.0	1755	12	AD157673 Human bre

22	36	60.0	1774	12	ADK43991 Plasmid J
23	36	60.0	1774	12	ADP32625 Human MUC
24	36	60.0	1799	12	AD023124 Human MUC
25	36	60.0	1803	12	AD157699 Human bre
26	36	60.0	1804	6	ABL67539 Thyroid c
27	36	60.0	1804	9	AAV56950 Human muc
28	36	60.0	1804	10	AD14719 Human bre
29	36	60.0	1804	12	ADP3294 Renal cel
30	36	60.0	1805	12	AD028642 Human MUC
31	36	60.0	1808	12	AD157706 Human bre
32	36	60.0	1818	12	ADP32633 Plasmid J
33	36	60.0	1823	6	AB235228 Human gen
34	36	60.0	1823	12	AD157707 Human bre
35	36	60.0	1835	12	ADP32631 Plasmid J
36	36	60.0	1874	12	AD157688 Human bre
37	36	60.0	1882	12	AD157677 Human bre
38	36	60.0	1918	12	AD157672 Human bre
39	36	60.0	1930	12	AD157678 Human bre
40	36	60.0	1945	12	AD157676 Human bre
41	36	60.0	1949	12	AD157698 Human bre
42	36	60.0	1953	12	AD157668 Human bre
43	36	60.0	2045	12	AD157701 Human bre
44	36	60.0	2049	12	AD157682 Human bre
45	36	60.0	2090	12	AD157705 Human bre

ALIGNMENTS

RESULT 1	AAV48324	standard; DNA; 60 BP.
ID	AAV48324	standard; DNA; 60 BP.
XX	AAV48324	
AC	AAV48324	
XX	20-NOV-1998	(first entry)
DT	20-NOV-1998	(first entry)
XX		
DE	Nucleotide sequence encoding MUC1 tandem repeat unit R9.	
XX	ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;	
KW	tumour; tumour-associated antigen.	
KW		
OS	Homo sapiens.	
XX		
XX	24-FEB-1998; 98WO-US003693.	
XX	24-FEB-1998; 98WO-US003693.	
XX	24-FEB-1997; 97US-0038253P.	
XX	24-FEB-1997; 97US-0038253P.	
XX	(THER-) THERION BIOLOGICS CORP.	
PA	(USSH) US DEPT HEALTH & HUMAN SERVICES.	
PA	(DAND) DANA-FABER CANCER INST INC.	
XX	Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;	
PI	WPI; 1998-467492/40.	
DR		
XX	New recombinant pox virus for tumour therapy - comprises DNA encoding an	
PT	immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat unit.	
PT		
XX	Disclosure; Page 11; 42pp; English.	
PS		
XX	The MUC1 tandem repeat unit AAV48324-V48325 were used to create an	
CC	immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus	
CC	(RPV). The RPV was used in a pharmaceutical composition also containing	
CC	an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The	
CC	recombinant pox virus therefore encodes an immunogenic MUC1 fragment that	
CC	does not undergo significant genetic deletion, thereby providing an	
CC	unexpectedly stable and immunogenic pox virus. They can be used to	
CC	prevent or treat tumours expressing MUC1 tumour-associated antigens	

```
XX
SQ Sequence 60 BP; 12 A; 26 C; 14 G; 8 T; 0 U; 0 Other;
Query Match 100.0%; Score 60; DB 2; Length 60;
Best Local Similarity 100.0%; Pred. No. 1.4e-10;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 GGTTCACGGCAGCTTCAGACACGAGTCTGACCCGACCCGCTCAGCTCCG 60
DB 1 GGTTCACGGCAGCTTCAGACACGAGTCTGACCCGACCCGCTCAGCTCCG 60
RESULT 2
ADK68635/c
ID ADK68635 standard; cDNA; 156 BP.
XX
AC ADK68635;
XX
DT 06-MAY-2004 (first entry)
XX
DE HSP65-MUC1 antigen CTL epitope related cDNA #3.
XX
KM Human breast cancer; Mycobacterium bovis; heat shock protein 65; BCG;
KW HSP65; cytotoxic T lymphocyte; CTL; MUC1; HSP65-MUC1 antigen CTL epitope;
XX ss.
XX Unidentified.
XX
XX CN1368384-A.
XX
XX 11-SEP-2002.
XX
XX 08-FEB-2001; 2001CN-00102614.
XX
XX 08-FEB-2001; 2001CN-00102614.
XX
XX (DIWE-) DIWEIHUAYU BIO TECHNOLOGY CO LTD BEIJING.
XX
XX Yu Y, Li H;
XX
XX WPI; 2003-854662/80.
XX
XX Genetically engineered vaccine of MUC-1 antigen for human breast cancer.
XX
XX Example 2; Page 4 (Disclosure); 14pp; Chinese.
XX
XX The invention relates to a method of preparation of a genetically
XX engineered vaccine for preventing and treating human breast cancer. The
XX method comprises fusing the coding gene of the Mycobacterium bovis heat
XX shock protein 65 (BCG HSP65) with the cytotoxic T lymphocyte (CTL)
XX epitope gene of MUC1 (antigen cell expressed by human breast cancer
XX cells) and the HSP65-MUC1 antigen CTL epitope is expressed in Escherichia
XX coli cells. This sequence represents DNA used in the method of the
XX invention.
XX
XX Sequence 156 BP; 37 A; 42 C; 58 G; 19 T; 0 U; 0 Other;
Query Match 68.0%; Score 40.8; DB 10; Length 156;
Best Local Similarity 80.0%; Pred. No. 0.00041;
Matches 48; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
OY 1 GGTTCACGGCAGCTTCAGACACGAGTCTGACCCGACCCGCTCAGCTCCG 60
DB 132 GGTTCACGGCAGCTTCAGACACGAGTCTGACCCGACCCGCTCAGCTCCG 73
RESULT 3
ADK68629
ID ADK68629 standard; cDNA; 1800 BP.
XX
AC ADK68629;
XX
DT 06-MAY-2004 (first entry)
```

```
XX
DE HSP65-MUC1 antigen CTL epitope related cDNA #1.
XX
XX Human breast cancer; Mycobacterium bovis; heat shock protein 65; BCG;
KW HSP65; cytotoxic T lymphocyte; CTL; MUC1; HSP65-MUC1 antigen CTL epitope;
XX gene; ss.
XX Unidentified.
XX
XX CN1368384-A.
XX
XX 11-SEP-2002.
XX
XX 08-FEB-2001; 2001CN-00102614.
XX
XX 08-FEB-2001; 2001CN-00102614.
XX
XX (DIWE-) DIWEIHUAYU BIO TECHNOLOGY CO LTD BEIJING.
XX
XX Yu Y, Li H;
XX
XX WPI; 2003-854662/80.
XX
XX P-PSDB; ADK68630.
XX
XX Genetically engineered vaccine of MUC-1 antigen for human breast cancer.
XX
XX Claim 1; SEQ ID NO 1; 14pp; Chinese.
XX
XX The invention relates to a method of preparation of a genetically
XX engineered vaccine for preventing and treating human breast cancer. The
XX method comprises fusing the coding gene of the Mycobacterium bovis heat
XX shock protein 65 (BCG HSP65) with the cytotoxic T lymphocyte (CTL)
XX epitope gene of MUC1 (antigen cell expressed by human breast cancer
XX cells) and the HSP65-MUC1 antigen CTL epitope is expressed in Escherichia
XX coli cells. This sequence represents DNA used in the method of the
XX invention.
XX
XX Sequence 1800 BP; 347 A; 561 C; 616 G; 276 T; 0 U; 0 Other;
Query Match 65.3%; Score 39.2; DB 10; Length 1800;
Best Local Similarity 78.3%; Pred. No. 0.002;
Matches 47; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
OY 1 GGTTCACGGCAGCTTCAGACACGAGTCTGACCCGACCCGCTCAGCTCCG 60
DB 1639 GGTTCACGGCAGCTTCAGACACGAGTCTGACCCGACCCGCTCAGCTCCG 1698
RESULT 4
AAV48318
ID AAV48318 standard; DNA; 60 BP.
XX
XX AAV48318;
XX
XX 20-NOV-1998 (first entry)
XX
XX Nucleotide sequence encoding MUC1 tandem repeat unit R3.
XX
XX ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
KW tumour; tumour-associated antigen.
XX
XX Homo sapiens.
XX
XX WO9837095-A2.
XX
XX 27-AUG-1998.
XX
XX 24-FEB-1998; 98WO-US003693.
XX
XX 24-FEB-1997; 97US-0038253P.
XX
XX (THER-) THERION BIOLOGICS CORP.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
```


PA (DAND) DANA FARBER CANCER INST INC.
XX
PI Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;
XX
DR WPI; 1998-467492/40.
XX
PT New recombinant pox virus for tumour therapy - comprises DNA encoding an
XX immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
XX
PS Disclosure; Page 11; 42pp; English.
XX
CC The MUC1 tandem repeat units AAV48317-V48325 were used to create an
CC immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
CC (RPV). The RPV was used in a pharmaceutical composition also containing
CC an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The
CC recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
CC does not undergo significant genetic deletion, thereby providing an
CC unexpectedly stable and immunogenic pox virus. They can be used to
CC prevent or treat tumours expressing MUC1 tumour-associated antigens
XX
SQ Sequence 60 BP; 7 A; 29 C; 19 G; 5 T; 0 U; 0 Other;
XX
Query Match 63.7%; Score 38.2; DB 2; Length 60;
Best Local Similarity 78.0%; Pred. No. 0.0026;
Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0
XX
QY 1 GATTCAACGGCACCTCTGACGACAGGAGTCACTGTGACCCGACCCCGTCCAGCTCC 59
Db 1 GGATCCACCGCGCCCGCTGCGGACCGAGTGAAGTCACTGTGCGGCCCGACACGCGCCCGCTCC 59
XX
RESULT 5
AAV48319
ID AAV48319 standard; DNA; 60 BP.
AC
XX AAV48319;
AC
XX 20-NOV-1998 (first entry)
XX
DT
XX
DE Nucleotide sequence encoding MUC1 tandem repeat unit R4.
XX
DE
XX
KW ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
KM tumour; tumour-associated antigen.
XX
XX Homo sapiens.
OS
XX
XX MO9637095-A2.
XX
XX 27-AUG-1998.
XX
XX 24-FEB-1998; 98WO-US003693.
XX
XX 24-FEB-1997; 97US-0038253P.
XX
XX (THER-) THERION BIOLOGICS CORP.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX (DAND) DANA FARBER CANCER INST INC.
XX
XX Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;
PI
XX WPI; 1998-467492/40.
XX
XX New recombinant pox virus for tumour therapy - comprises DNA encoding an
XX immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
XX
XX Disclosure; Page 11; 42pp; English.
XX
XX The MUC1 tandem repeat units AAV48317-V48325 were used to create an
XX immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
XX (RPV). The RPV was used in a pharmaceutical composition also containing
XX an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The
XX recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
XX does not undergo significant genetic deletion, thereby providing an

CC	unexpectably stable and immunogenic pox virus. They can be used to
CC	prevent or treat tumours expressing MUC1 tumour-associated antigens
XX	
50	Sequence 60 BP; 10 A; 23 C; 13 G; 14 T; 0 U; 0 Other;
Qy	Query Match 61.0%; Score 36.6; DB 2; Length 60;
Db	Best Local Similarity 76.3%; Pred. No. 0.009;
XX	Matches 45; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
1	GGTTCAACGGACCTCCGAGCACACGAGTCAGTCTGACACCCGACACCGCTCCAGCTCC 59
1	GGGTCAACAGCTCTCTCCGCTCATGGGGTTACTTCTCTCCAGATACCTCCGCGAGCTCC 59
RESULT 6	
ID	AAV48316 standard; cDNA; 60 BP.
XX	AAV48316;
AC	
XX	AAV48316;
DT	20-NOV-1998 (first entry)
XX	
DE	Nucleotide sequence encoding MUC1 tandem repeat unit.
XX	
KM	BS; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
KW	tumour; tumour-associated antigen.
XX	
OS	Homo sapiens.
XX	
PH	Key
FT	Location/Qualifiers
FT	1..60
CDS	/*tag= a
FT	/product= "MUC1 tandem repeat unit"
PN	MO9837095-A2.
XX	
PD	27-AUG-1998.
XX	
PF	24-FEB-1998; 98MO-US003693.
XX	
PR	24-FEB-1997; 97US-0038253P.
XX	
PA	(THER-) THERION BIOLOGICS CORP.
XX	(USSH) US DEPT HEALTH & HUMAN SERVICES.
PA	(DAND) DANA FARBEN CANCER INST INC.
PI	Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;
XX	
DR	WPI; 1998-467492/40.
XX	
DR	P-PSDB; AAW71229.
XX	
PT	New recombinant pox virus for tumour therapy - comprises DNA encoding an
XX	immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
XX	
PS	Example 1; Page 20; 42pp; English.
XX	
CC	The MUC1 tandem repeat unit was used to create an immunogenic mini-MUC1
CC	fragment for inclusion in a recombinant pox virus (RPV). The RPV was used
CC	in a pharmaceutical composition also containing an immunomodulator to
CC	generate MUC1 specific cytotoxic T-lymphocytes. The recombinant pox virus
CC	therefore encodes an immunogenic MUC1 fragment that does not undergo
CC	significant genetic deletion, thereby providing an unexpectedly stable
CC	and immunogenic pox virus. They can be used to prevent or treat tumours
CC	expressing MUC1 tumour-associated antigens
XX	
SQ	Sequence 60 BP; 7 A; 33 C; 16 G; 4 T; 0 U; 0 Other;
XX	
Query Match	60.0%; Score 36; DB 2; Length 60;
Best Local Similarity	75.0%; Pred. No. 0.014;
Matches	45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
1	GGTTCAACGGACCTCCGAGCACACGAGTCAGTCTGACACCCGACACCGCTCCAGCTCCG 60
1	GGTTCAACGGACCTCCGAGCACACGAGTCAGTCTGACACCCGACACCGCTCCAGCTCCG 60

Db 1 GGCTCCAGCGCCCCCGACGCCAGGTGTACCTCGGCCCGGACACAGCGCCGCCCG 60

RESULT 7
AAV48326
ID AAV48326 standard; cDNA; 78 BP.
XX
AC AAV48326;
XX
DT 20-NOV-1998 (first entry)
XX
DE Nucleotide sequence encoding MUC1 tandem repeat unit a.
XX
KM ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
XX
OS tumour; tumour-associated antigen.
XX
FH Homo sapiens.
XX
FH Key Location/Qualifiers
FT 1..78
FT /*tag= a
FT /product= "MUC1 tandem repeat unit"
XX
PN MO9837095-A2.
XX
PD 27-AUG-1998.
XX
PF 24-FEB-1998; 98WO-US003693.
XX
PR 24-FEB-1997; 97US-0038253P.
XX
PA (THER-) THERION BIOLOGICS CORP.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PA (DAND) DANA-FARBER CANCER INST INC.
XX
PI Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;
XX
DR WPI: 1998-467492/40.
XX
DR P-PSDB; AAW77230.
XX
PT New recombinant pox virus for tumour therapy - comprises DNA encoding an
XX
PT immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
XX
PS Example 1; Page 20; 42pp; English.
XX
CC The MUC1 tandem repeat units AAV48326-V48328 were used to create an
XX
CC immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
XX
CC (RPV). The RPV was used in a pharmaceutical composition also containing
XX
CC an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The
XX
CC recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
XX
CC does not undergo significant genetic deletion, thereby providing an
XX
CC unexpectedly stable and immunogenic pox virus. They can be used to
XX
CC prevent or treat tumours expressing MUC1 tumour-associated antigens
XX
SQ Sequence 78 BP; 9 A; 42 C; 22 G; 5 T; 0 U; 0 Other;

Query Match 60.0%; Score 36; DB 2; Length 78;
Best Local Similarity 75.0%; Pred. No. 0.015;
Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

OY 1 GGTTCAACGGCAGCTCCAGACACGAGTACGTCTGACCCGACACCGCTCAGCTCCG 60
DB 1 GGCTCCAGCGCCCCCGACGCCAGGTGTACCTCGGCCCGGACACAGCGCCGCCCG 60

RESULT 8
AAN90579/c
ID AAN90579 standard; cDNA; 309 BP.
XX
AC AAN90579;
XX
DT 27-AUG-2003 (revised)
XX
DT 25-MAR-2003 (revised)

DT 04-DEC-1989 (first entry)
XX
XX PDF9.3 cDNA insert.
XX
XX PDF9.3; human DF3 breast carcinoma-associated antigen epitope.
XX
OS Human MCF-7 breast carcinoma cells; 'ATCC HTB22'.
XX
PN MO8907107-A.
XX
PD 10-AUG-1989.
XX
PF 29-JAN-1988; 88US-00149831.
XX
PR 29-JAN-1988; 88US-00149831.
XX
PA (DANA-) DANA-FARBER CANCER.
XX
PI Kufe DW;
XX
DR WPI: 1989-248989/34.
XX
DR P-PSDB; AAP91045, AAP91053, AAP91054, AAP90146.
XX
PT Recombinant polypeptide(s) - contains DF 3 breast carcinoma antigen
XX
PT epitope and useful as assay reagents, and encoding DNA sequences.
XX
PS Claim 1; Fig 4; 31pp; English.
XX
CC The sequence encodes a carbohydrate-free polypeptide contg. DF3 breast
XX
CC carcinoma antigen epitope. Useful as a competitive binding assay reagent
XX
CC and improves diagnosis. The cDNA consists of nearly identical 60 BP
XX
CC tandem repeats which are 85% GC-rich. See also AAP91045, AAP91046,
XX
CC AAP91053, AAP91054. (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 309 BP; 21 A; 92 C; 160 G; 36 T; 0 U; 0 Other;

Query Match 60.0%; Score 36; DB 1; Length 309;
Best Local Similarity 75.0%; Pred. No. 0.018;
Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

OY 1 GGTTCAACGGCAGCTCCAGACACGAGTACGTCTGACCCGACACCGCTCAGCTCCG 60
DB 266 GGCTCCAGCGCCCCCGACGCCAGGTGTACCTCGGCCCGGACACAGCGCCGCCCG 207

RESULT 9
AD157712
ID AD157712 standard; cDNA; 1194 BP.
XX
AC AD157712;
XX
DT 22-APR-2004 (first entry)
XX
DE Human breast specific nucleic acid (BSNA) #83.
XX
XX Human breast specific nucleic acid; BSNA; gene; ss; metastasis;
XX
KM Human breast cancer; cytostatic.
XX
OS Homo sapiens.
XX
PN WO2003106648-A2.
XX
PD 24-DEC-2003.
XX
PF 16-JUN-2003; 2003WO-US018934.
XX
PR 14-JUN-2002; 2002US-0389327P.
XX
PA (DIAD-) DIADEXUS INC.
XX
PI Salceda S, Macina RA, Turner LR, Sun Y, Liu C;
XX

DR WPI; 2004-082185/08.
DR P-PSDB; ADI57782.
XX
PT Novel isolated polypeptide comprising breast specific protein sequences,
PT useful for diagnosing or monitoring presence and metastases of breast
cancer in patient.
XX
PS Claim 1; SEQ ID NO 83; 370pp; English.
XX
CC The invention relates to human breast specific nucleic acids (BSNA) and
CC the breast specific proteins (BSP) they encode. The nucleic acids are
CC useful for determining the presence of a BSNA in a sample which involves
CC contacting the sample with a BSNA under conditions in which the BSNA will
CC selectively hybridise to a BSNA in the sample, and detecting the
CC hybridisation. The nucleic acids are useful for determining the presence
CC of a BSP in a sample which involves contacting the sample with suitable
CC reagent under conditions in which the reagent will selectively interact
CC with the BSP, and detecting the interaction of the reagent with a BSP in
CC the sample. The nucleic acids and proteins are useful for diagnosing or
CC monitoring the presence and metastases of breast cancer in a patient,
CC which involves determining an amount of nucleic acid or protein and
CC comparing the determined amount of nucleic acid or protein in the sample
CC of the patient to the amount of a breast specific marker in a normal
CC control, where a difference in the determined amount in the sample
CC compared to the amount in the control is associated with the presence of
CC breast cancer. The sequences are useful for treating a patient with
CC breast cancer, involving administering a composition consisting of a BSNA
CC or a BSP to a patient, where the administration induces an immune
CC response against the breast cancer cell expressing the BSNA or BSP. This
CC sequence represents a human BSNA of the invention.
SQ Sequence 1194 BP; 243 A; 438 C; 275 G; 238 T; 0 U; 0 Other;
Query Match 60.0%; Score 36; DB 12; Length 1194;
Best Local Similarity 75.0%; Pred. No. 0.022;
Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 1 GGTTCACGGCAGCTCAGACACGAGTCTGACACCGGACACCGCTCCAGCTCCG 60
DB 596 GGCTCCACCGCCGCCGCCGACCGGTGTACCTCGGCCCGGACACGAGCGGCCCG 655
RESULT 10
ADI57693
ID ADI57693 standard; cDNA; 1378 BP.
XX
XX ADI57693;
AC
XX
DT 22-APR-2004 (first entry)
XX
DE Human breast specific nucleic acid (BSNA) #64.
XX
XX Human; breast specific nucleic acid; BSNA; gene; ss; metastasis;
KW breast cancer; cytosolic.
XX
XX Homo sapiens.
OS
XX
XX WO2003106648-A2.
PN
XX
XX 24-DEC-2003.
PD
XX
XX 16-JUN-2003; 2003WO-US018934.
PF
XX
XX 14-JUN-2002; 2002US-0389327P.
PR
XX
XX (DIAD-) DIADEXUS INC.
PA
XX
XX Salceda S, Macina RA, Turner LR, Sun Y, Liu C;
PI
XX
XX WPI; 2004-082185/08.
DR
XX
XX P-PSDB; ADI57765.
PT
XX
XX Novel isolated polypeptide comprising breast specific protein sequences.

PT useful for diagnosing or monitoring presence and metastases of breast
PT cancer in patient.
XX
XX
PS Claim 1; SEQ ID NO 64; 370pp; English.
XX
XX
CC The invention relates to human breast specific nucleic acids (BSNA) and
CC the breast specific proteins (BSP) they encode. The nucleic acids are
CC useful for determining the presence of a BSNA in a sample which involves
CC contacting the sample with a BSNA under conditions in which the BSNA will
CC selectively hybridise to a BSNA in the sample, and detecting the
CC hybridisation. The nucleic acids are useful for determining the presence
CC of a BSP in a sample which involves contacting the sample with suitable
CC reagent under conditions in which the reagent will selectively interact
CC with the BSP, and detecting the interaction of the reagent with a BSP in
CC the sample. The nucleic acids and proteins are useful for diagnosing or
CC monitoring the presence and metastases of breast cancer in a patient,
CC which involves determining an amount of nucleic acid or protein and
CC comparing the determined amount of nucleic acid or protein in the sample
CC of the patient to the amount of a breast specific marker in a normal
CC control, where a difference in the determined amount in the sample
CC compared to the amount in the control is associated with the presence of
CC breast cancer. The sequences are useful for treating a patient with
CC breast cancer, involving administering a composition consisting of a BSNA
CC or a BSP to a patient, where the administration induces an immune
CC response against the breast cancer cell expressing the BSNA or BSP. This
CC sequence represents a human BSNA of the invention.
SQ Sequence 1378 BP; 295 A; 489 C; 316 G; 278 T; 0 U; 0 Other;
Query Match 60.0%; Score 36; DB 12; Length 1378;
Best Local Similarity 75.0%; Pred. No. 0.022;
Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 1 GGTTCACGGCAGCTCAGACACGAGTCTGACACCGGACACCGCTCCAGCTCCG 60
DB 596 GGCTCCACCGCCGCCGCCGACCGGTGTACCTCGGCCCGGACACGAGCGGCCCG 655
RESULT 11
ADO23180/C
ID ADO23180 standard; RNA; 1424 BP.
XX
XX ADO23180;
AC
XX
DT 12-AUG-2004 (first entry)
XX
DE Antisense human MUC1 mucin glycoprotein RNA (coding sequence) SeqID 75.
XX
XX human; MUC1; mucin glycoprotein; cancer; chemotherapeutic; MUC1/BCD; ss;
KW antisense.
XX
XX Homo sapiens.
OS
XX
XX WO2004044160-A2.
PN
XX
XX 27-MAY-2004.
PD
XX
XX 12-NOV-2003; 2003WO-US035848.
PF
XX
XX 13-NOV-2002; 2002US-00293391.
PR
XX
XX 29-MAY-2003; 2003US-00447839.
PA
XX
XX (DAND) DANA FARBER CANCER INST INC.
PA
XX
XX (ILEX-) ILEX PROD INC.
PI
XX
XX Kufe DW, Kharbanda S, Weltman SD;
PI
XX
XX WPI; 2004-420304/39.
DR
XX
XX Double-stranded RNA complex useful for inhibiting proliferation of cancer
PT cell expressing MUC1 mucin glycoprotein, comprises first and second RNA
PT sequences.

PS Disclosure; SEQ ID NO 75; 112pp; English.

XX
CC This invention relates to novel modulators of the human MUC1 mucin
CC glycoprotein for use in cancer therapeutics, where MUC1 is a protein that
CC acts to inhibit the apoptotic response to genotoxic stress caused by
CC chemotherapeutic agents. In particular, it refers to modulators of the
CC MUC1 extracellular domain (MUC1/ECD). The method refers to using double-
CC stranded RNA complexes as MUC1 interference RNA compositions such that
CC MUC1 expression is inhibited, which in turn inhibits cancer cell
CC proliferation. The present invention describes screening assays to
CC identify compounds that inhibit the binding of various MUC1 ligands such
CC as neurigin 2, as well as agonists, antagonists and antibodies thereof.
CC Furthermore, it provides MUC1 antisense, small interfering RNA (siRNA)
CC and small molecules in combination with chemotherapeutic agents that are
CC useful in the field of cancer therapy. This polynucleotide sequence is
CC the antisense human MUC1 RNA of the invention.

SQ Sequence 1424 BP; 309 A; 325 C; 495 G; 0 T; 295 U; 0 Other;

Query Match 60.0%; Score 36; DB 12; Length 1424;
Best Local Similarity 75.0%; Pred. No. 0.022;

Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 GGTTACAGGCGACCTCCAGCACGAGTCACTGTGACCCGACACCGGTCCAGTCCG 60
DB 1041 GGCTCCACCGCCCCCGCCAGCCACGCTGTACCTCGGCCCGACACAGGCGGCCCG 982

RESULT 12
ABL60159

ID ABL60159 standard; cDNA; 1428 BP.

XX ABL60159;

AC 22-JUN-2002 (first entry)

DE Human MUC1 encoding cDNA SEQ ID NO 2.

XX Human mucin 1; MUC1; transmembrane protein; SNP; cancer; cytostatic;
KW single nucleotide polymorphism; haplotyping; genotyping; drug;
XX antiinflammatory; gene; ss.

OS Homo sapiens.

XX Key Location/Qualifiers
FH 1..1428

FT CDS /*tag= a
FT /*product= "MUC1"

FT variation /*tag= b
FT /*tag= b
FT /*tag= b

FT /standard name= "Single nucleotide polymorphism"
FT /note= "SNP allelic variation results in Val substituted
FT by Met at position 337 of the MUC1 protein (AB077476)"
XX

PN W0200226765-A2.

PD 04-APR-2002.

XX 25-SEP-2001; 2001WO-US030151.

XX 28-SEP-2000; 2000US-0236113P.

XX (GENA-) GENAISSANCE PHARM INC.

PI Chew A, Koshy B;

DR WPI; 2002-405042/43.

XX P-PsDB; AB077476.

PT New genetic variants of mucin 1, Transmembrane gene, useful in studying
PT expression and function of protein encoded by the gene and for screening
PT drugs to treat diseases e.g. cancer.
XX

PS Claim 23; Fig 2; 75pp; English.

XX
CC The invention relates to a polynucleotide (ABL60158, ABL60159) encoding
CC mucin 1/MUC1 (AB077476), Transmembrane isogene. The invention describes
CC novel genetic variants of the MUC1 gene. The invention is useful for
CC haplotyping/genotyping the MUC1 gene in an individual and identifying an
CC association between a trait and at least one of the haplotypes or
CC haplotype pairs of MUC1 gene. MUC1 is useful for studying the expression
CC and function of MUC1 gene. MUC1 is useful for use in screening for
CC candidate drugs to treat diseases related to MUC1 activity and in
CC studying the effect of the variation on the biological activity of MUC1
CC as well as on the binding affinity of candidate drugs targeting MUC1 for
CC the treatment of e.g. cancer. MUC1 is further used by the pharmaceutical
CC research scientist to validate MUC1 as a candidate target for and in
CC design of clinical trials of candidate drugs for, treating a specific
CC condition drugs or disease predicted to be associated with MUC1 activity.
CC MUC1 antibodies are useful in a variety of diagnostic and prognostic
CC formats and therapeutic methods
XX

SQ Sequence 1428 BP; 296 A; 496 C; 327 G; 309 T; 0 U; 0 Other;

Query Match 60.0%; Score 36; DB 6; Length 1428;
Best Local Similarity 75.0%; Pred. No. 0.022;

Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 GGTTACAGGCGACCTCCAGCACGAGTCACTGTGACCCGACACCGGTCCAGTCCG 60
DB 385 GGCTCCACCGCCCCCGCCAGCCACGCTGTACCTCGGCCCGACACAGGCGGCCCG 444

RESULT 13

ID ADO23125 standard; RNA; 1428 BP.

XX ADO23125;

DE 12-AUG-2004 (first entry)

XX Human MUC1 mucin glycoprotein RNA (coding sequence) SeqID 20.

XX human; MUC1; mucin glycoprotein; cancer; chemotherapeutic; MUC1/BCD; ss.

XX Homo sapiens.

XX W02004044160-A2.

PD 27-MAY-2004.

XX 12-NOV-2003; 2003WO-US035848.

XX 13-NOV-2002; 2002US-00293391.

XX 29-MAY-2003; 2003US-00447839.

XX (DAND) DANA FARBER CANCER INST INC.

PI (IDEX-) ILEX PROD INC.

DR Kufe DW, Kharbanda S, Weltman SD;

XX WPI; 2004-420304/39.

PS Claim 2; SEQ ID NO 20; 112pp; English.

XX This invention relates to novel modulators of the human MUC1 mucin
CC glycoprotein for use in cancer therapeutics, where MUC1 is a protein that
CC acts to inhibit the apoptotic response to genotoxic stress caused by
CC chemotherapeutic agents. In particular, it refers to modulators of the
CC MUC1 extracellular domain (MUC1/BCD). The method refers to using double-
CC stranded RNA complexes as MUC1 interference RNA compositions such that
CC MUC1 expression is inhibited, which in turn inhibits cancer cell

CC proliferation. The present invention describes screening assays to
CC identify compounds that inhibit the binding of various MUC1 ligands such
CC as neurigin 2, as well as agonists, antagonists and antibodies thereof.
CC Furthermore, it provides MUC1 antisense, small interfering RNA (siRNA)
CC and small molecules in combination with chemotherapeutic agents that are
CC useful in the field of cancer therapy. This polynucleotide sequence is
CC the human MUC1 RNA of the invention.

XX Sequence 1428 BP; 296 A; 496 C; 327 G; 0 T; 309 U; 0 Other;

Query Match 60.0%; Score 36; DB 12; Length 1428;
Best Local Similarity 70.0%; Pred. No. 0.022;
Matches 42; Conservative 3; Mismatches 15; Indels 0; Gaps 0;

QY 1 GGTTCACGGCAGCTCCAGACACGAGTGTGACCCGACCCGCTCCAGCTCCG 60
DB 385 GGCTCCACGGCCCCCGCCAGCCAGGUGUACCCGCCCCGACACAGCGCGCCCG 444

RESULT 14

ADP32627 standard; DNA; 1457 BP.

AC ADF32627;

XX 26-FEB-2004 (first entry)

DE Plasmid JMW358 MUC-1 nucleotide sequence.

XX MUC-1 antigen; immune response; MUC-1; variable number of tandem repeat;

KM VNTR; repeat unit; tumour; metastasis; cytostatic; vaccine; gene therapy;

XX gene; db.

OS Synthetic.

XX MO200310060-A2.

XX 04-DEC-2003.

PF 23-MAY-2003; 2003WO-EP005594.

PR 24-MAY-2002; 2002GB-00012046.

PA (GLAXO) GLAXO GROUP LTD.

PI Burden N, Ellis JH, Hamblin PA;

DR WPI; 2004-042811/04.

XX New nucleic acid molecule encoding a MUC-1 antigen, useful for preparing

PT a composition for treating or preventing tumors or metastases.

PS Example; Fig 3; 66pp; English.

XX The present invention describes a nucleic acid molecule which encodes a
CC MUC-1 antigen. The nucleic acid is capable of raising an immune response
CC in vivo, has reduced susceptibility to recombination than full-length MUC
CC -1 and comprises between 1 and 15 variable number of tandem repeats
CC (VNTR) perfect repeat units. Also described: (1) a plasmid comprising the
CC DNA molecule; (2) a protein encoded by the nucleic acid; (3) a
CC pharmaceutical composition comprising the nucleic acid, plasmid or
CC protein and an excipient, diluent or carrier; and (4) a method of
CC treating or preventing tumors or metastases. A MUC1 antigen has
CC cytostatic activity, and can be used in vaccines, and in gene therapy.
CC The nucleic acid is useful for preparing a composition for treating or
CC preventing tumors or metastases. The present sequence is used in the
CC exemplification of the present invention.

XX Sequence 1457 BP; 302 A; 506 C; 334 G; 315 T; 0 U; 0 Other;

Query Match 60.0%; Score 36; DB 12; Length 1457;
Best Local Similarity 75.0%; Pred. No. 0.022;
Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 GGTTCACGGCAGCTCCAGACACGAGTGTGACCCGACCCGCTCCAGCTCCG 60
DB 403 GGCTCCACGGCCCCCGCCAGCCAGGTGTACCTCGGCCCGACACAGCGCGCCCG 462

RESULT 15

AAV48329 standard; cDNA; 1527 BP.

XX AAV48329;

AC 20-NOV-1998 (first entry)

XX MiniMUC1 gene.

DE MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;

KM tumour; tumour-associated antigen.

XX Homo sapiens.

OS Key Location/Qualifiers

FT CDS 1..1527 /tag= a

XX /product= "MiniMUC1 protein"

XX MO9837095-A2.

XX 27-AUG-1998.

XX 24-FEB-1998; 98WO-US003693.

XX 24-FEB-1997; 97US-0038253P.

XX (THER-) THERION BIOLOGICS CORP.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

PA (DAND) DANA FARBER CANCER INST INC.

XX Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;

PI WPI; 1998-467492/40.

DR P-PSDB; AAW77233.

XX New recombinant pox virus for tumour therapy - comprises DNA encoding an

PT immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.

PS Example 1; Page 21-22; 42pp; English.

XX The immunogenic mini-MUC1 fragment was created from MUC1 tandem repeat
CC units for inclusion in a recombinant pox virus (RPV). The RPV was used in
CC a pharmaceutical composition also containing an immunomodulator to
CC generate MUC1 specific cytotoxic T-lymphocytes. The recombinant pox virus
CC therefore encodes an immunogenic MUC1 fragment that does not undergo
CC significant genetic deletion, thereby providing an unexpectedly stable
CC and immunogenic pox virus. They can be used to prevent or treat tumours
CC expressing MUC1 tumour-associated antigens

XX Sequence 1527 BP; 296 A; 573 C; 351 G; 307 T; 0 U; 0 Other;

Query Match 60.0%; Score 36; DB 2; Length 1527;
Best Local Similarity 75.0%; Pred. No. 0.023;
Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 GGTTCACGGCAGCTCCAGACACGAGTGTGACCCGACCCGCTCCAGCTCCG 60
DB 226 GGCTCCACGGCCCCCGCCAGCCAGGTGTACCTCGGCCCGACACAGCGCGCCCG 285

Search completed: January 15, 2005, 20:36:11
Job time : 170.5 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2005, 17:40:52 : Search time 722.8 Seconds
(without alignments)
3925.542 Million cell updates/sec

Title: US-10-057-136-11

Perfect score: 60

Sequence: 1 ggtctcaacgcacacccctccagc.....ccgacacccgtccagctccg 60

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 2364849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_hgt:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pac:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_srb:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	2297	6	BD272907 A recombi
2	36	60.0	120	6	AX192396 Sequence
3	36	60.0	120	6	BD000571 Human pol
4	36	60.0	1414	12	AF423031 Synthetic
5	36	60.0	1455	6	CQ715242 Sequence
6	36	60.0	1457	6	AX959914 Sequence
7	36	60.0	1774	6	AX959914 Sequence
8	36	60.0	1774	6	AX959912 Sequence
9	36	60.0	1804	6	AR492318 Sequence
10	36	60.0	1804	6	AX335367 Sequence
11	36	60.0	1804	9	HUMMUCAB
12	36	60.0	1834	12	AF423030
13	36	60.0	1835	6	AX959918 Sequence
14	36	60.0	2135	6	AX959916 Sequence
15	36	60.0	4139	6	CQ834017 Sequence
16	36	60.0	4139	6	AX334899 Sequence
17	36	60.0	4139	6	AX335372 Sequence
18	36	60.0	4139	6	AX336712 Sequence
19	36	60.0	4139	6	AX409474 Sequence

20	36	60.0	4139	6	AX440481	AX440481 Sequence
21	36	60.0	4139	9	HUMEPANMU	J05582 Human pancr
22	36	60.0	7188	9	AY463543	AY463543 Homo sapi
23	36	60.0	8181	6	AX406624	AX406624 Sequence
24	36	60.0	8181	6	HUMPEM	Me1170 Human polyim
25	36	60.0	8186	6	AR492320	AR492320 Sequence
26	36	60.0	133525	9	AL713999	AL713999 Human DNA
27	34.4	57.3	120	6	AX192397	AX192397 Sequence
28	34.4	57.3	120	6	BD000572	BD000572 Human pol
29	34.4	57.3	180	6	E08763	E08763 CDNA encodi
30	34.4	57.3	572	6	AR492319	AR492319 Sequence
31	34.4	57.3	572	6	HUMDF3AA	M31823 Human breas
32	34.4	57.3	1721	6	CQ711290	CQ711290 Sequence
33	34.4	57.3	1721	6	AR492306	AR492306 Sequence
34	34.4	57.3	1721	6	AX335860	AX335860 Sequence
35	34.4	57.3	1721	6	AX440427	AX440427 Sequence
36	34.4	57.3	1721	6	AX587588	AX587588 Sequence
37	34.4	57.3	1721	6	HSTERYMA	X52229 Human mRNA
38	34.4	57.3	1971	6	AX963157	AX963157 Sequence
39	34.4	57.3	2037	6	AX963159	AX963159 Sequence
40	34.4	57.3	2238	9	HSSRTA	X52228 Human mRNA
41	34.4	57.3	2751	9	AF176947	AF176947 Macaca mu
42	33.6	56.0	491	9	HUMEPISIB1	M32739 Human epist
43	33.6	56.0	518	6	AR492402	AR492402 Sequence
44	33.6	56.0	518	9	HUMEPISIA1	M32738 Human epist
45	33.6	56.0	1320	6	E08764	E08764 CDNA encodi

ALIGNMENTS

RESULT 1	BD272907	2297 bp	DNA	linear	PAT 17-JUL-2003
LOCUS	BD272907				
DEFINITION	A recombinant vector expressing multiple constitutatory molecules and uses thereof.				
ACCESSION	BD272907				
VERSION	BD272907.1	GI:33082675			
KEYWORDS	JP 2002531133-A/1.				
SOURCE	synthetic construct				
ORGANISM	artificial sequences.				
REFERENCE	1 (bases 1 to 2297)				
AUTHORS	Schlom,J., Hodge,J. and Panicali,D.				
TITLE	A recombinant vector expressing multiple constitutatory molecules and uses thereof				
JOURNAL	Patent: JP 2002531133-A 1 24-SEP-2002;				
COMMENT	THE UNITED STATES OF AMERICA, THERION BIOLOGICS CORP				
OS	Artificial Sequence				
PN	JP 2002531133-A/1				
PD	24-SEP-2002				
PF	12-NOV-1999 JP 2000586927				
PR	09-DEC-1998 US 60/111582				
PI	JEFFREY SCHLOM, JAMES HODGE, DENNIS PANICALI				
PC	C12N15/02, A61K35/12, A61K35/74, A61K35/76, A61K38/00, A61K39/00,				
PC	A61K39/12,				
PC	A61K39/125, A61K39/21, A61K39/235, A61K39/245, A61K39/275, A61K39/				
PC	29, A61K48/00,				
PC	A61P1/04, A61P31/04, A61P31/10, A61P31/12, A61P35/00, A61P37/02, PC				
PC	A61P37/06,				
PC	C12N1/15, C12N1/19, C12N1/21, C12N5/10, C12N7/00, C12Q1/02, G01N33/				
PC	53, C12N15/00,				
PC	C12N5/00, A61K37/02				
CC	Description of Artificial Sequence: VECTOR SEQUENCE FH				
FT	Location/Qualifiers				
FT	source 1..2297				
FEATURES	Location/Qualifiers				
source	1..2297				
ORIGIN	/organism="synthetic construct"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:32630"				

Query Match
Best Local Similarity 100.0%; Score 60; DB 6; Length 2297;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCACGGCAGCCTCCAGCAGGAGTCACTGTCACCCGACACCCGCTCCAGTCCG 60
DB 826 GGTTCACGGCAGCCTCCAGCAGGAGTCACTGTCACCCGACACCCGCTCCAGTCCG 885

RESULT 2
LOCUS AX192396 120 bp DNA linear PAT 15-AUG-2001
DEFINITION Sequence 2 from Patent EP1103623.
ACCESSION AX192396
VERSION AX192396.1 GI:15210363
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 Taylor-Papadimitriou, J., Burchell, J. and Gendler, S.
AUTHORS Human mucin core protein: nucleic acid probes, peptide fragments
TITLE and antibodies thereco, and uses thereof in diagnostic and
JOURNAL therapeutic methods
FEATURES Patent: EP 1103623-A 2 30-MAY-2001;
source IMPERIAL CANCER RESEARCH TECHNOLOGY LIMITED (GB)
Location/Qualifiers
1..120
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match
Best Local Similarity 75.0%; Score 36; DB 6; Length 120;
Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 GGTTCACGGCAGCCTCCAGCAGGAGTCACTGTCACCCGACACCCGCTCCAGTCCG 60
DB 87 GGCTCCACGGCCCCCCCCAGCCAGCGGTGTCACTCGGCCCGACACAGCGCGCCCG 28

RESULT 3
LOCUS BD000571 120 bp DNA linear PAT 31-JAN-2002
DEFINITION Human polymorphic epithelial mucin core protein and nucleic acid
ACCESSION BD000571 encoding the protein.
VERSION BD000571.1 GI:18623684
KEYWORDS UP 2000333675-A/2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 120)
AUTHORS Papadimitriou, J., Jendora, S. and Bachle, J.
TITLE Human polymorphic epithelial mucin core protein and nucleic acid
JOURNAL Patent: JP 2000333675-A 2 05-DEC-2000;
COMMENT IMPERIAL CANCER RESEARCH TECHNOLOGY LTD
OS Homo sapiens (human)
PN JP 2000333675-A/2
PD 05-DEC-2000
PR 26-APR-2000 JP 2000125724
PR 07-JAN-1987 GB 8700269, 07-JAN-1987 GB 8700279 PR
22-APR-1987 US 041306, 09-NOV-1987 GB 8726172 PI JOYCE
TAYLOR PAPADIMITRIOU, SANDRA JENDORA, JOY BACHLE PC
C12N15/02, A61K38/00, A61K39/395, A61K49/00, A61P35/00, PC
C07K16/44, C12N5/10, C12P21/08 // (C12N15/02, C12R1:91), (C12N5/10,
PC C12R1:91),

PC C12N15/00, A61K37/02, C12N5/00, (C12N15/00, C12R1:91), (C12N5/00,
PC C12R1:91)
CC
FH
FT
FT
Key Location/Qualifiers
source 1..120
/organism="Homo sapiens (human)"
Location/Qualifiers
1..120
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match
Best Local Similarity 60.0%; Score 36; DB 6; Length 120;
Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 GGTTCACGGCAGCCTCCAGCAGGAGTCACTGTCACCCGACACCCGCTCCAGTCCG 60
DB 87 GGCTCCACGGCCCCCCCCAGCCAGCGGTGTCACTCGGCCCGACACAGCGCGCCCG 28

RESULT 4
LOCUS AF423031 1414 bp mRNA linear SYN 10-JUL-2003
DEFINITION Synthetic construct Homo sapiens mucin variant MUC1-CT58 (MUC1)
ACCESSION AF423031 mRNA, complete cds; alternatively spliced.
VERSION AF423031
KEYWORDS AF423031.1 GI:19338621
SOURCE
ORGANISM
synthetic construct
synthetic construct
artificial sequences.

REFERENCE 1 (bases 1 to 1414)
AUTHORS Hinojosa-Kurtzberg, A.M., Johansson, M.E., Madsen, C.S., Hansson, G.C.
TITLE Hinojosa-Kurtzberg, A.M., Johansson, M.E., Madsen, C.S., Hansson, G.C.
JOURNAL Novel MUC1 splice variants contribute to mucin overexpression in
FEATURES Am. J. Physiol. Gastrointest. Liver Physiol. 284 (5), G853-G862
(2003)
MEDLINE 22570517
PUBMED 12529261
REFERENCE 2 (bases 1 to 1414)
AUTHORS Hinojosa-Kurtzberg, A.M. and Gendler, S.J.
TITLE Direct Submission
JOURNAL Submitted (21-SEP-2001) Biochemistry, Mayo Clinic, 13400 E. Shea
Boulevard, Scottsdale, AZ 85259, USA
Location/Qualifiers
1..1414
/organism="synthetic construct"
/mol_type="mRNA"
/db_xref="taxon:32630"
/note="Homo sapiens gene in transgenic Mus musculus
C57Bl/6; isolated from intestinal mucosa"
1..1414
/gene="MUC1"
1..1386
/gene="MUC1"
/note="alternatively spliced; contains exon 6b resulting
in variant carboxy-terminal domain; lacks sites for
beta-catenin and Grib2 interactions; derived from Homo
sapiens"
/codon_start=1
/transl_table=1
/product="mucin variant MUC1-CT58"
/protein_id="AA186735.1"
/db_xref="GI:19338622"
/translation="MTPTGQSPFFLLTLTVLVGSGHASTPGKETSATORS
VPSSTENAVSMISVLSHSPGSGSTGQGVTLAPATBPASGSAATWQDVTSV
VTRPALGSTTPPAHDVTSAPDNKPAVGSAPPAHVTAPDRPAGTAPPAAGVTS
APDNRPALGSTAPVNVTSASGASGASTLVNHTGTSARATTTASSTPFSIPSH
SDPTTASHTKSTDSSTHSTHSTVPLTSPNSHSTSPQSTGVSFFLSFHSIQFNS
SLDPSDVTYQGLQRISEMFQITKGGFGLSLNFKRPSPSVVQTLAEREETINV

CDS

gene

CDS

ORIGIN
Query Match 60.0%; Score 36; DB 12; Length 1414;
Best Local Similarity 75.0%; Pred. No. 0.45;
Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
HDEVTFQNYKTEAASRYNLITSDVSVDPFPFSAQSGAVGMIALLVLCVLA
LAIVYLALAVCCCRKXNYGQDLIFPARDYHMSRYPTVHTGRVVPSSSTRSRYE
KEMRVDRRELA"

QY 1 GGTTCACGGCAGCTCCAGACACGAGTCACTGTCACCCGACCCGCTCCAGCTCCG 60
DB 385 GGCTCCACGGCCCCCGACGCCAGGTGTCACTCGGCCGACACACGAGCCGCCCG 444

RESULT 5
LOCUS AX959914 1455 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 1176 from Patent WO2068579.
ACCESSION CO715242
VERSION CO715242.1 GI:42276099
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
TITLE Kite, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
Patent: WO 02068579-A 1176 06-SEP-2002;
PB Corporation (NY) (US)
Location/Qualifiers
1. .1455
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 60.0%; Score 36; DB 6; Length 1455;
Best Local Similarity 75.0%; Pred. No. 0.45;
Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
1 GGTTCACGGCAGCTCCAGACACGAGTCACTGTCACCCGACCCGCTCCAGCTCCG 60
DB 412 GGCTCCACGGCCCCCGACGCCAGGTGTCACTCGGCCGACACACGAGCCGCCCG 471

QY 1 GGTTCACGGCAGCTCCAGACACGAGTCACTGTCACCCGACCCGCTCCAGCTCCG 60
DB 412 GGCTCCACGGCCCCCGACGCCAGGTGTCACTCGGCCGACACACGAGCCGCCCG 471

RESULT 6
LOCUS AX959914 1457 bp DNA linear PAT 14-JAN-2004
DEFINITION Sequence 19 from Patent WO03100060.
ACCESSION AX959914
VERSION AX959914.1 GI:40880143
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Burden, N.G., Ellis, J.H. and Hamblin, P.A.
TITLE Muc-1 antigen with reduced number of vnter repeat units
JOURNAL Patent: WO 03100060-A 19 04-DEC-2003;
GLAXO GROUP LIMITED (GB)
Location/Qualifiers
1. .1457
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 60.0%; Score 36; DB 6; Length 1457;
Best Local Similarity 75.0%; Pred. No. 0.45;
Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
1 GGTTCACGGCAGCTCCAGACACGAGTCACTGTCACCCGACCCGCTCCAGCTCCG 60
DB 462 GGCTCCACGGCCCCCGACGCCAGGTGTCACTCGGCCGACACACGAGCCGCCCG 521

ORIGIN
Query Match 60.0%; Score 36; DB 6; Length 1457;
Best Local Similarity 75.0%; Pred. No. 0.45;
Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
1 GGTTCACGGCAGCTCCAGACACGAGTCACTGTCACCCGACCCGCTCCAGCTCCG 60
DB 462 GGCTCCACGGCCCCCGACGCCAGGTGTCACTCGGCCGACACACGAGCCGCCCG 521

Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 1 GGTTCACGGCAGCTCCAGACACGAGTCACTGTCACCCGACCCGCTCCAGCTCCG 60
DB 403 GGCTCCACGGCCCCCGACGCCAGGTGTCACTCGGCCGACACACGAGCCGCCCG 462

RESULT 7
LOCUS AX959684 1774 bp DNA linear PAT 14-JAN-2004
DEFINITION Sequence 10 from Patent WO03099193.
ACCESSION AX959684
VERSION AX959684.1 GI:40880030
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Burden, N.G., Ellis, J.H. and Hamblin, P.A.
TITLE Vaccines
JOURNAL Patent: WO 03099193-A 10 04-DEC-2003;
GLAXO GROUP LIMITED (GB)
Location/Qualifiers
1. .1774
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 60.0%; Score 36; DB 6; Length 1774;
Best Local Similarity 75.0%; Pred. No. 0.44;
Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
1 GGTTCACGGCAGCTCCAGACACGAGTCACTGTCACCCGACCCGCTCCAGCTCCG 60
DB 462 GGCTCCACGGCCCCCGACGCCAGGTGTCACTCGGCCGACACACGAGCCGCCCG 521

QY 1 GGTTCACGGCAGCTCCAGACACGAGTCACTGTCACCCGACCCGCTCCAGCTCCG 60
DB 462 GGCTCCACGGCCCCCGACGCCAGGTGTCACTCGGCCGACACACGAGCCGCCCG 521

RESULT 8
LOCUS AX959912 1774 bp DNA linear PAT 14-JAN-2004
DEFINITION Sequence 17 from Patent WO03100060.
ACCESSION AX959912
VERSION AX959912.1 GI:40880142
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Burden, N.G., Ellis, J.H. and Hamblin, P.A.
TITLE Muc-1 antigen with reduced number of vnter repeat units
JOURNAL Patent: WO 03100060-A 17 04-DEC-2003;
GLAXO GROUP LIMITED (GB)
Location/Qualifiers
1. .1774
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 60.0%; Score 36; DB 6; Length 1774;
Best Local Similarity 75.0%; Pred. No. 0.44;
Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
1 GGTTCACGGCAGCTCCAGACACGAGTCACTGTCACCCGACCCGCTCCAGCTCCG 60
DB 462 GGCTCCACGGCCCCCGACGCCAGGTGTCACTCGGCCGACACACGAGCCGCCCG 521

ORIGIN
Query Match 60.0%; Score 36; DB 6; Length 1774;
Best Local Similarity 75.0%; Pred. No. 0.44;
Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
1 GGTTCACGGCAGCTCCAGACACGAGTCACTGTCACCCGACCCGCTCCAGCTCCG 60
DB 462 GGCTCCACGGCCCCCGACGCCAGGTGTCACTCGGCCGACACACGAGCCGCCCG 521

LOCUS AR492318 1804 bp DNA linear PAT 15-MAY-2004
 DEFINITION Sequence 17 from patent US 6716627.
 ACCESSION AR492318
 VERSION AR492318.1 GI:47260892
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE Unclassified.
 AUTHORS Dobie, K.W.
 TITLE 1 (bases 1 to 1804)
 JOURNAL Antisense modulation of mucin 1, transmembrane expression
 FEATURES Patent: US 6716627-A 17 06-APR-2004;
 Location/Qualifiers
 1..1804
 /organism="unknown"
 /mol_type="genomic DNA"

ORIGIN

Query Match 60.0%; Score 36; DB 6; Length 1804;
 Best Local Similarity 75.0%; Pred. No. 0.44;
 Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 GGTTCACGGCAGCCTCCAGACACGAGTCACTGTGACCCGACACCCGTCAGCTCCG 60
 457 GGCTCCACCGCCCCCCCCAGCCAGGTGTCACTCGGCCCGGACACACGAGCCGCCCG 516

RESULT 10
 LOCUS AX335367 1804 bp DNA linear PAT 09-JAN-2002
 DEFINITION Sequence 5876 from Patent WO0194629.
 ACCESSION AX335367
 VERSION AX335367.1 GI:18126086
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
 Horrigan, S., Soppet, D.R. and Weaver, Z.
 TITLE Cancer gene determination and therapeutic screening using signature
 JOURNAL Patent: WO 0194629-A 5876 13-DEC-2001;
 Avalon Pharmaceuticals (US)
 FEATURES Location/Qualifiers
 1..1804
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN

Query Match 60.0%; Score 36; DB 6; Length 1804;
 Best Local Similarity 75.0%; Pred. No. 0.44;
 Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 GGTTCACGGCAGCCTCCAGACACGAGTCACTGTGACCCGACACCCGTCAGCTCCG 60
 457 GGCTCCACCGCCCCCCCCAGCCAGGTGTCACTCGGCCCGGACACACGAGCCGCCCG 516

RESULT 11
 LOCUS HUMNUCAB 1804 bp mRNA linear PRI 07-JAN-1995
 DEFINITION Human polymorphic epithelial mucin (PEM) mRNA, complete cde.
 ACCESSION J05581
 VERSION J05581.1 GI:188869
 KEYWORDS polymorphic epithelial mucin.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1804)

AUTHORS Gendler, S.J., Lancaster, C.A., Taylor-Papadimitriou, J., Dhill, T.,
 Peat, N., Burchell, J., Pemberton, L., Lalani, S.N. and Wilson, D.
 TITLE Molecular cloning and expression of human tumor-associated
 JOURNAL polymorphic epithelial mucin
 MEDLINE J. Biol. Chem. 265 (25), 15286-15293 (1990)
 PUBMED 1697589
 COMMENT Original source text: Homo sapiens adult adenocarcinoma cDNA to
 mRNA.
 Draft entry and computer-readable sequence for [J. Biol. Chem.
 (1990) In press] kindly submitted
 by S.J.Gendler, 26-JUN-1990.

FEATURES
 source
 1..1804
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /map="XG26.3-q27.1"
 /cell_line="BT20"
 /tissue_type="adenocarcinoma"
 /dev_stage="adult"
 1..1804
 /gene="MCNAA"
 <1..1804
 /gene="MCNAA"
 /product="PEM mRNA"
 73..1500
 /gene="MCNAA"
 /product="PEM mRNA"
 /codon_start=1
 /product="polymorphic epithelial mucin"
 /protein_id="AA59876.1"
 /db_xref="GI:188870"
 /translation="MTPTGSPFLLITLVITVYTGSGHASTPFGKEKTSATORS
 VPSSTKNAVSMSTSVLSHSGSSSTTGGQDVTLAPATPAGSATAWQDVTSVP
 VTRPALGSTTPAHVDVTSAPDKPAPGSTAPPAHGVTSAPTRPAGSTAPPAHGVTS
 APNRPALGSTAPPAHGVTSAPDKPAPGSTAPPAHGVTSAPTRPAGSTAPPAHGVTS
 SDPTTTLASHSTKTDASTHSTHSTVPLTSSHSTSPOLSTGVSPFISPHISNQPMS
 SLDEPSTDYVQELORDISEMFLQIKQGFGLSNIKRPGSVVQVLTIAFREBTIV
 HDVETQPNQVKTBAASRYNLITSDVSVDFPFSAGSAGAPGKGIALLVLCVLA
 LAIVYLLALAVOCRRKNYGGDLIPPADYTHPMSVPTTHTRKIVYPSSTDRSPYE
 KYAGGSGSSLSYTNPAVAATSNL"
 73..1135
 /gene="MCNAA"
 1136..1497
 /gene="MCNAA"
 /product="polymorphic epithelial mucin"
 1783..1788
 /gene="MCNAA"

polya_signal

mat_peptide

sig_peptide

ORIGIN

Query Match 60.0%; Score 36; DB 9; Length 1804;
 Best Local Similarity 75.0%; Pred. No. 0.44;
 Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 GGTTCACGGCAGCCTCCAGACACGAGTCACTGTGACCCGACACCCGTCAGCTCCG 60
 457 GGCTCCACCGCCCCCCCCAGCCAGGTGTCACTCGGCCCGGACACACGAGCCGCCCG 516

RESULT 12
 LOCUS AF423030 1834 bp mRNA linear SYN 10-JUL-2003
 DEFINITION Synthetic construct Homo sapiens mucin variant MUC-1CT80 (MUC1)
 ACCESSION AF423030
 VERSION AF423030
 KEYWORDS
 SOURCE AF423030.1 GI:19338619
 ORGANISM
 REFERENCE 1 (bases 1 to 1834)
 AUTHORS Hinojosa-Kurtzberg, A.M., Johanson, M.E., Madsen, C.S., Hansson, G.C.
 and Gendler, S.J.

TITLE Novel MUC1 splice variants contribute to mucin overexpression in
CFTR-deficient mice
JOURNAL Am. J. Physiol. Gastrointest. Liver Physiol. 284 (5), G853-G862
(2003)
MEDLINE 22570517
PUBMED 12529261
REFERENCE 2 (bases 1 to 1834)
AUTHORS Hinojosa-Kurtzberg, A.M. and Gendler, S.J.
TITLE Direct Submission
JOURNAL Submitted (21-SEP-2001) Biochemistry, Mayo Clinic, 13400 E. Shea
Boulevard, Scottsdale, AZ 85259, USA
Location/Qualifiers
source 1. 1834
/organism="synthetic construct"
/mol_type="mRNA"
/db_xref="taxon:32630"
/note="Homo sapiens gene in transgenic Mus musculus
C57BL/6; isolated from intestinal mucosa"
1. 1834
/gene="MUC1"
1. 1452
/gene="MUC1"
/note="alternatively spliced; contains exon 6a resulting
in variant hydrophilic tail of carboxy-terminal domain;
lacks sites for beta-catenin and Grb2 interactions;
derived from Homo sapiens"
/codon_start=1
/transl_table=11
/product="mucin variant MUC-1CT80"
/protein_id="AAL86734.1"
/db_xref="GI:19338620"
/translation="MPTGQSPPELILLTLTVYVSGHASTPGEKETSATQSS
VPSSTKNAVSMSTSVLSHSGSGSSSTTGQDYLAPTEPASGAATGQDVSVP
VTRPALGTTTPAHVDTSAPDNKPAFGSTAPPAHGTSAADTRPAGSTAPPAHGT
ADNRPALGSTAPPAHVTASGASGASGASGASGASGASGASGASGASGASGASG
SDPTPLASTKTDASTHSTHSTHSTHSTHSTHSTHSTHSTHSTHSTHSTHSTHSTH
SLEDSTDYDGELOKDISSEMPLOIKOGFGLGSLNKKRPGSVVVLTAPEGTIN
HVEIQPNQYKTEBAASRYVLTISDVSVDPPEPASGAGVPGMGLIALIVVCVIVA
LAIVYLIALAVQCCKRKNYGQDIDIPARTYHPMSERYTHYHGRVPPSSIDRSRYE
KRLPPTGGKORVWLGDSEGTWKTQAMRK"

gene
CDS
1. 1834
/gene="MUC1"
1. 1452
/gene="MUC1"
/note="alternatively spliced; contains exon 6a resulting
in variant hydrophilic tail of carboxy-terminal domain;
lacks sites for beta-catenin and Grb2 interactions;
derived from Homo sapiens"
/codon_start=1
/transl_table=11
/product="mucin variant MUC-1CT80"
/protein_id="AAL86734.1"
/db_xref="GI:19338620"
/translation="MPTGQSPPELILLTLTVYVSGHASTPGEKETSATQSS
VPSSTKNAVSMSTSVLSHSGSGSSSTTGQDYLAPTEPASGAATGQDVSVP
VTRPALGTTTPAHVDTSAPDNKPAFGSTAPPAHGTSAADTRPAGSTAPPAHGT
ADNRPALGSTAPPAHVTASGASGASGASGASGASGASGASGASGASGASGASG
SDPTPLASTKTDASTHSTHSTHSTHSTHSTHSTHSTHSTHSTHSTHSTHSTHSTH
SLEDSTDYDGELOKDISSEMPLOIKOGFGLGSLNKKRPGSVVVLTAPEGTIN
HVEIQPNQYKTEBAASRYVLTISDVSVDPPEPASGAGVPGMGLIALIVVCVIVA
LAIVYLIALAVQCCKRKNYGQDIDIPARTYHPMSERYTHYHGRVPPSSIDRSRYE
KRLPPTGGKORVWLGDSEGTWKTQAMRK"

ORIGIN

Query Match 60.0%; Score 36; DB 12; Length 1834;
Best Local Similarity 75.0%; Pred. No. 0.44;
Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 1 GGTTCAACGGCAGCTTCACAGCAGGAGTCAGCTGTGACACCCGCTCAGCTCCG 60
Db 365 GGCTCACCGCCCCCGCCAGCCAGGTGTCACTCGGCCCGGACACGAGCGGCCCG 444

RESULT 13
AX959918 1835 bp DNA linear PAT 14-JAN-2004
LOCUS
DEFINITION Sequence 23 from Patent WO03100060.
ACCESSION AX959918
VERSION AX959918.1 GI:40880145
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 Burden, N.G., Ellis, J.H. and Hamblin, P.A.
AUTHORS Muc-1 antigen with reduced number of vnter repeat units
TITLE Patent: WO 03100060-A 23 04-DEC-2003;
JOURNAL GLAXO GROUP LIMITED (GB)
FEATURES Location/Qualifiers
source 1. 1835
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 60.0%; Score 36; DB 6; Length 1835;
Best Local Similarity 75.0%; Pred. No. 0.44;
Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 1 GGTTCAACGGCAGCTTCACAGCAGGAGTCAGCTGTGACACCCGCTCAGCTCCG 60
Db 462 GGCTCACCGCCCCCGCCAGCCAGGTGTCACTCGGCCCGGACACGAGCGGCCCG 521

RESULT 14
AX959916 2135 bp DNA linear PAT 14-JAN-2004
LOCUS
DEFINITION Sequence 21 from Patent WO03100060.
ACCESSION AX959916
VERSION AX959916.1 GI:40880144
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 Burden, N.G., Ellis, J.H. and Hamblin, P.A.
AUTHORS Muc-1 antigen with reduced number of vnter repeat units
TITLE Patent: WO 03100060-A 21 04-DEC-2003;
JOURNAL GLAXO GROUP LIMITED (GB)
FEATURES Location/Qualifiers
source 1. 2135
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 60.0%; Score 36; DB 6; Length 2135;
Best Local Similarity 75.0%; Pred. No. 0.44;
Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 1 GGTTCAACGGCAGCTTCACAGCAGGAGTCAGCTGTGACACCCGCTCAGCTCCG 60
Db 462 GGCTCACCGCCCCCGCCAGCCAGGTGTCACTCGGCCCGGACACGAGCGGCCCG 521

RESULT 15
CQ834017 4139 bp DNA linear PAT 29-JUL-2004
LOCUS
DEFINITION Sequence 53 from Patent EPI439393.
ACCESSION CQ834017
VERSION CQ834017.1 GI:50833622
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 Astle, J.H., Boardman, L.A., Bugart, L.J., Burgess, C.C., Catino, T.J.,
AUTHORS Diwedi, P., Huntress, M., Johnson, K.A., Lewis, M.E., Matmonis, P.J.,
Myerow, S.H., Brown-Shimer, S.L., Thiagalingam, A., Thibodeau, S.N. and
Molino, G.A.
TITLE Detection methods using T1MP 1 for colon cancer diagnosis
JOURNAL Patent: EP 1439393-A 53 21-JUL-2004;
Bayer Healthcare LLC (US); MAYO FOUNDATION FOR MEDICAL EDUCATION
AND RESEARCH (US)

FEATURES Location/Qualifiers
source 1. 4139
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 60.0%; Score 36; DB 6; Length 4139;
Best Local Similarity 75.0%; Pred. No. 0.41;
Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Sun Jan 16 16:35:01 2005

us-10-057-136-11.rge

Page 6

QY 1 GGTTCAAGGCACCTTCAGACACGAGTACAGTGTGACCCGACACCCGTCAGCTCG 60
DB 458 GGCTTCACCGCCCCCCAGCCACGGTGTCACTTCGAGCCCGACACCGAGCGGGCCCG 517

Search completed: January 15, 2005, 22:36:54
Job time : 722.8 secs

Db 1 GGGTCGACTGCCCTCCGGCGCATGTGTGACTTCACTCCTTGACACAGGCCAGCCCCA 60

RESULT 2

US-10-406-317-41
; Sequence 41, Application US/10406317
; Publication No. US20040019195A1
; GENERAL INFORMATION:
; APPLICANT: Schlom, Jeffrey;
; APPLICANT: Hodges, James;
; APPLICANT: Panicali, Dennis
; TITLE OF INVENTION: A recombinant vector expressing multiple constitutively
; FILE REFERENCE: 38163-0189
; CURRENT APPLICATION NUMBER: US/10/406,317
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: US/09/856,988
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: PCT/US99/26866
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/111,582
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41
; LENGTH: 2297
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: VECTOR
US-10-406-317-41

Query Match
Best Local Similarity 100.0%; Score 60; DB 16; Length 2297;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGTCGACTGCCCTCCGGCGCATGTGTGACTTCACTCCTTGACACAGGCCAGCCCCA 60
Db 766 GGGTCGACTGCCCTCCGGCGCATGTGTGACTTCACTCCTTGACACAGGCCAGCCCCA 825

RESULT 3

US-10-057-136-2
; Sequence 2, Application US/10057136
; Publication No. US20030021770A1
; GENERAL INFORMATION:
; APPLICANT: SCHLOM, JEFFREY
; APPLICANT: KANTOR, JUDITH
; APPLICANT: KUFE, DONALD
; APPLICANT: PANICALI, DENNIS
; APPLICANT: GRITZ, LINDA
; TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MCM1
; FILE REFERENCE: 700953/47113C
; CURRENT APPLICATION NUMBER: US/10/057,136
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/366,670
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: PCT/US98/03693
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/038,253
; PRIOR FILING DATE: 1997-02-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-136-2

Query Match
Best Local Similarity 78.0%; Score 38.2; DB 14; Length 60;
Matches 42; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 1 GGGTCGACTGCCCTCCGGCGCATGTGTGACTTCACTCCTTGACACAGGCCAGCCCC 59
Db 1 GGCTCCACCGCCCGCCCGCAGCCGATGTGACTTCCGCGGACACAGGCCCGCCCC 59

RESULT 4

US-10-447-839A-75/c
; Sequence 75, Application US/10447839A
; Publication No. US20040018181A1
; GENERAL INFORMATION:
; APPLICANT: Kufe, Donald W.
; APPLICANT: Kharbanda, Sunder
; APPLICANT: Weitman, Steven D.
; TITLE OF INVENTION: MCM1 INTERFERENCE RNA COMPOSITIONS AND METHODS DERIVED THEREFROM
; FILE REFERENCE: 1000.05.009
; CURRENT APPLICATION NUMBER: US/10/447,839A
; PRIOR FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: 10/293,391
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 09/951,938
; PRIOR FILING DATE: 2001-09-11
; PRIOR APPLICATION NUMBER: 60/231,841
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 75
; LENGTH: 1424
; TYPE: RNA
; ORGANISM: ARTIFICIAL
; FEATURE:
; OTHER INFORMATION: Synthesized Sequence
US-10-447-839A-75

Query Match
Best Local Similarity 63.7%; Score 38.2; DB 16; Length 1424;
Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 1 GGGTCGACTGCCCTCCGGCGCATGTGTGACTTCACTCCTTGACACAGGCCAGCCCC 59
Db 1041 GGCTCCACCGCCCGCCCGCAGCCGATGTGACTTCCGCGGACACAGGCCCGCCCC 983

RESULT 5

US-10-447-839A-20
; Sequence 20, Application US/10447839A
; Publication No. US20040018181A1
; GENERAL INFORMATION:
; APPLICANT: Kufe, Donald W.
; APPLICANT: Kharbanda, Sunder
; APPLICANT: Weitman, Steven D.
; TITLE OF INVENTION: MCM1 INTERFERENCE RNA COMPOSITIONS AND METHODS DERIVED THEREFROM
; FILE REFERENCE: 1000.05.009
; CURRENT APPLICATION NUMBER: US/10/447,839A
; PRIOR FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: 10/293,391
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 09/951,938
; PRIOR FILING DATE: 2001-09-11
; PRIOR APPLICATION NUMBER: 60/231,841
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 20
; LENGTH: 1428
; TYPE: RNA
; ORGANISM: RNA
US-10-447-839A-20

Query Match
Best Local Similarity 63.7%; Score 38.2; DB 16; Length 1428;
Matches 42; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

```
QY      1 GGGTCGACTGCCCTCCCGGCAGTGTGTGAACCTTAGCCTTGACAACAGGCCACGCC 59
         || : | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     385 GGCUCCACCGCCCCCGCCAGCCCAAGUGUACAUUGGGCCC CGAACCAGGCCGCC 443
```

```

? RESULT 6
? US-10-057-136-19
? Sequence 19, Application US/10057136
? Publication No. US20030021770A1
? GENERAL INFORMATION:
? APPLICANT: SCHLOW, JEFFREY
? APPLICANT: KANTOR, JUDITH
? APPLICANT: KUFE, DONALD
? APPLICANT: PANICALI, DENNIS
? APPLICANT: GRITZ, LINDA
? TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR
? TITLE OF INVENTION: TUMOR-ASSOCIATED ANTIGEN
? FILE REFERENCE: 700953/47113C
? CURRENT APPLICATION NUMBER: US/10/057,136
? CURRENT FILING DATE: 2002-01-25
? PRIOR APPLICATION NUMBER: 09/366,670
? PRIOR FILING DATE: 1999-08-03
? PRIOR APPLICATION NUMBER: PCT/US98/03693
? PRIOR FILING DATE: 1998-02-24
? PRIOR APPLICATION NUMBER: 60/038,253
? PRIOR FILING DATE: 1997-02-24
? NUMBER OF SEQ ID NOS: 20
? SOFTWARE: PatentIn Ver. 2.1
? SEQ ID NO 19
? LENGTH: 1527
? TYPE: DNA
? ORGANISM: Homo sapiens
? US-10-057-136-19

```

Query Match	73.7%	Score 38.2	DB 14	Length 1527
Best Local Similarity	78.0%	Pred. No. 0.00015		
Matches	46	Conservative 0	Mismatches 11	Indels 0
			Gaps	0
Qy	1	GGGTCGACTGCGCCCTCCGCGCGCATGCTGTGTGACCTCACTCTCTGACACACAGGCGACGGCC	59	
Db	226	GGGCTTCACGCGCGCCCGCCGACCCACGATGTGTGACCTTGTGCGCGCGGACACACAGGCGCGGCC	284	

```

RESULT 7
US-10-447-839A-19
; Sequence 19, Application US/10447839A
; Publication No. US20040018181A1
; GENERAL INFORMATION:
; APPLICANT: Kufe, Donald W.
; APPLICANT: Kharbada, Surrender
; APPLICANT: Weltman, Steven D.
; TITLE OF INVENTION: MOC1 INTERFERENCE RNA COMPOSITIONS AND METHODS DERIVED THEREFROM
; FILE REFERENCE: 1000.05.009
; CURRENT APPLICATION NUMBER: US/10/447,839A
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: 10/293,391
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 09/951,938
; PRIOR FILING DATE: 2001-09-11
; PRIOR APPLICATION NUMBER: 60/231,841
; PRIOR FILING DATE: 2000-09-11
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 19
; LENGTH: 1799
; TYPE: RNA
; ORGANISM: RNA
; US-10-447-839A-19

```

Query Match	63.7%	Score 38.2;	DB 16;	Length 1799;
Best Local Similarity	71.2%	Pred. No. 0.00015;		
Matches 42;	Conservative 4;	Mismatches 13;	Indels 0;	Gaps 0

```

Oy      1 GGGTGCATGCCCTTCGGCGCATGGTGTGACTCAGTCTCTACACAAAGGCCAGCCCC 59
        |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      453 GGCUCACGCGCCCCCGCAGCCACGGGUGACCCUCCGCCCCGSHACCAAGCGCCCC 511

```

```

1 RESULT 8
2 US-09-964-824A-573
3 ; Sequence 573, Application US/09964824A
4 ; Patent No. US20020102531A1
5 ;
6 ; GENERAL INFORMATION:
7 ; APPLICANT: Horrigan, Stephen
8 ; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
9 ; TITLE OF INVENTION: Sets
10 ; FILE REFERENCE: 689230-73
11 ;
12 ; CURRENT APPLICATION NUMBER: US/09/964,824A
13 ; CURRENT FILING DATE: 2001-09-27
14 ;
15 ; PRIOR APPLICATION NUMBER: US/60/236,033
16 ; PRIOR FILING DATE: 2000-09-28
17 ;
18 ; PRIOR APPLICATION NUMBER: US/60/236,032
19 ; PRIOR FILING DATE: 2000-09-28
20 ;
21 ; PRIOR APPLICATION NUMBER: US/60/236,028
22 ; PRIOR FILING DATE: 2000-09-28
23 ;
24 ; NUMBER OF SEQ ID NOS: 583
25 ; SOFTWARE: PatentIn version 3.0
26 ;
27 ; SEQ ID NO 573
28 ; LENGTH: 1804
29 ; TYPE: DNA
30 ;
31 ; ORGANISM: Homo sapiens
32 ; US-09-964-824A-573

```

Query Match 63.7% ; Score 38.2 ; DB 9 ; Length 1804 ;
 Best Local Similarity 78.0% ; Pred. No. 0.00015 ;
 Matches 46 ; Conservative 0 ; Mismatches 13 ; Indels 0 ; Gaps 0 ;

```

RESULT 9
US-10-029-517-17
: Sequence 17, Application US//10029517
: Publication No. US20030148969A1
GENERAL INFORMATION:
APPLICANT: Kenneth W. Dobie
APPLICANT: Susan J. Myers
: TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
: FILE REFERENCE: RTS-0352
CURRENT APPLICATION NUMBER: US//10/029,517
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ. ID NOS: 107
SEQ. ID NO 17
LENGTH: 1804
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (73) ... (1500)
US-10-029-517-17

```

Query Match	63.7%	Score 38.2;	DB 15;	Length 1804;
Best Local Similarity	78.0%	Pred. No. 0.00015;		
Matches 46;	Conservative 0;	Mismatches 13;	Indels 0;	Gaps 0

Qy 1 GGGTTCAGCTGCCCTTCGGCGCATGGTGTGACCTCAGCTCTGTACACAGGCCAGCCCC 59
Db 457 GGCTTCACGCGCCCCCGACGCCACGGGTGACCTCGGCCCCGAGACACAGGCCGCGCCCC 515

RESULT 10
US-10-717-597-30
; Sequence 30, Application US/10717597

```
Publication No. US20040110221A1
GENERAL INFORMATION:
APPLICANT: Wyeck
APPLICANT: Burczynski, Michael E.
APPLICANT: Twine, Natalie C.
APPLICANT: Dornier, Andrew J.
APPLICANT: Trepichio, William L.
APPLICANT: Slonim, Donna K.
APPLICANT: Stever, Jennifer A.
TITLE OF INVENTION: METHODS FOR DIAGNOSING RCC AND OTHER SOLID TUMORS
FILE REFERENCE: AM1010801
CURRENT APPLICATION NUMBER: US/10/717,597
PRIOR FILING DATE: 2003-11-21
PRIOR APPLICATION NUMBER: US 60/459,782
PRIOR FILING DATE: 2003-04-03
PRIOR APPLICATION NUMBER: US 60/427,982
PRIOR FILING DATE: 2002-11-21
NUMBER OF SEQ ID NOS: 4904
SOFTWARE: PatentIn version 3.2
SEQ ID NO 30
LENGTH: 1804
TYPE: DNA
ORGANISM: Homo sapiens
US-10-717-597-30
```

```
Query Match      63.7%; Score 38.2; DB 17; Length 1804;
Best Local Similarity 78.0%; Pred. No. 0.00015;
Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
```

```
QY 1 GGGTCGACTGCCCCCTCCGGCGCATGTGTGACCTCCTGTGACACAGGCCAGCCCC 59
DB 457 GGCTCCACCGCCCCCCCCCAGCCACGCTGTCACTCGCCCGGACACAGGCCAGCCCC 515
```

```
RESULT 11
US-10-775-920-84
Sequence 84, Application US/10775920
Publication No. US20040175744A1
GENERAL INFORMATION:
APPLICANT: Merzen Ltd
TITLE OF INVENTION: POLYNUCLEOTIDE SEQUENCES AND CORRESPONDING ENCODED POLYPEPTIDES
TITLE OF INVENTION: OF PARTICULAR SECRETED AND MEMBRANE-BOUND PROTEINS OVEREXPRESSED
FILE REFERENCE: IN CERTAIN CANCERS
CURRENT APPLICATION NUMBER: US/10/775,920
CURRENT FILING DATE: 2004-02-10
PRIOR APPLICATION NUMBER: US 60/447,900
PRIOR FILING DATE: 2003-02-13
NUMBER OF SEQ ID NOS: 385
SOFTWARE: PatentIn version 3.2
SEQ ID NO 84
LENGTH: 1804
TYPE: DNA
ORGANISM: Homo sapiens
US-10-775-920-84
```

```
Query Match      63.7%; Score 38.2; DB 17; Length 1804;
Best Local Similarity 78.0%; Pred. No. 0.00015;
Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
```

```
QY 1 GGGTCGACTGCCCCCTCCGGCGCATGTGTGACCTCCTGTGACACAGGCCAGCCCC 59
DB 457 GGCTCCACCGCCCCCCCCCAGCCACGCTGTCACTCGCCCGGACACAGGCCAGCCCC 515
```

```
RESULT 12
US-10-101-510-339
Sequence 339, Application US/10101510
Publication No. US20030148295A1
GENERAL INFORMATION:
APPLICANT: WAN, JACKSON
APPLICANT: WANG, YIXIN
TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
```

```
FILE REFERENCE: 15117.0012
CURRENT APPLICATION NUMBER: US/10/101,510
CURRENT FILING DATE: 2002-03-20
PRIOR APPLICATION NUMBER: 60/276,947
PRIOR FILING DATE: 2001-03-20
NUMBER OF SEQ ID NOS: 805
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 339
LENGTH: 1823
TYPE: DNA
ORGANISM: Homo sapiens
US-10-101-510-339
```

```
Query Match      63.7%; Score 38.2; DB 15; Length 1823;
Best Local Similarity 78.0%; Pred. No. 0.00015;
Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
```

```
QY 1 GGGTCGACTGCCCCCTCCGGCGCATGTGTGACCTCCTGTGACACAGGCCAGCCCC 59
DB 457 GGCTCCACCGCCCCCCCCCAGCCACGCTGTCACTCGCCCGGACACAGGCCAGCCCC 515
```

```
RESULT 13
US-09-964-824A-105
Sequence 105, Application US/09964824A
Patent No. US20020102531A1
GENERAL INFORMATION:
APPLICANT: Horrigan, Stephen
TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signat
FILE REFERENCE: 689290-73
CURRENT APPLICATION NUMBER: US/09/964,824A
CURRENT FILING DATE: 2001-09-27
PRIOR APPLICATION NUMBER: US/60/236,033
PRIOR FILING DATE: 2000-09-28
PRIOR APPLICATION NUMBER: US/60/236,032
PRIOR FILING DATE: 2000-09-28
PRIOR APPLICATION NUMBER: US/60/236,028
NUMBER OF SEQ ID NOS: 583
SOFTWARE: PatentIn version 3.0
SEQ ID NO 105
LENGTH: 4139
TYPE: DNA
ORGANISM: Homo sapiens
US-09-964-824A-105
```

```
Query Match      63.7%; Score 38.2; DB 9; Length 4139;
Best Local Similarity 78.0%; Pred. No. 0.00015;
Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
```

```
QY 1 GGGTCGACTGCCCCCTCCGGCGCATGTGTGACCTCCTGTGACACAGGCCAGCCCC 59
DB 458 GGCTCCACCGCCCCCCCCCAGCCACGCTGTCACTCGCCCGGACACAGGCCAGCCCC 516
```

```
RESULT 14
US-09-964-824A-578
Sequence 578, Application US/09964824A
Patent No. US20020102531A1
GENERAL INFORMATION:
APPLICANT: Horrigan, Stephen
TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
FILE REFERENCE: 689290-73
CURRENT APPLICATION NUMBER: US/09/964,824A
CURRENT FILING DATE: 2001-09-27
PRIOR APPLICATION NUMBER: US/60/236,033
PRIOR FILING DATE: 2000-09-28
PRIOR APPLICATION NUMBER: US/60/236,032
PRIOR FILING DATE: 2000-09-28
PRIOR APPLICATION NUMBER: US/60/236,028
PRIOR FILING DATE: 2000-09-28
```


NUMBER OF SEQ ID NOS: 583
SOFTWARE: Patentin version 3.0
SEQ ID NO: 578
LENGTH: 4139
TYPE: DNA
ORGANISM: Homo sapiens
US-09-964-824A-578

Query Match 63.7%; Score 38.2; DB 9; Length 4139;
Best Local Similarity 78.0%; Pred. No. 0.00015;
Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 GGGTCGACTGCCCTCGGCGCATGTGTGACCTCAGCTCTTGACACAAAGCCAGCCCC 59
DB 458 GGCTCCACCGCCCCCGACGCCACGAGTGTACCTCGCCCCGAGACACAGCCCGGCCCC 516

RESULT 15

US-09-864-864-334
Sequence 334, Application US/09864864
Patent No. US2020102679A1
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan L.
APPLICANT: Dillon, Davin G.
APPLICANT: Secrist, Heather
APPLICANT: Lodes, Michael J.
APPLICANT: Algate, Paul A.
APPLICANT: Fling, Steve P.
APPLICANT: Mannion, Jane
APPLICANT: Benson, Darin R.
APPLICANT: Carter, Darick
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
FILE REFERENCE: 210121.523
CURRENT APPLICATION NUMBER: US/09/864,864
CURRENT FILING DATE: 2001-05-23
NUMBER OF SEQ ID NOS: 341
SOFTWARE: Corixa Invention Disclosure Database
SEQ ID NO: 334
LENGTH: 4139
TYPE: DNA
ORGANISM: Homo sapiens
US-09-864-864-334

Query Match 63.7%; Score 38.2; DB 9; Length 4139;
Best Local Similarity 78.0%; Pred. No. 0.00015;
Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 GGGTCGACTGCCCTCGGCGCATGTGTGACCTCAGCTCTTGACACAAAGCCAGCCCC 59
DB 458 GGCTCCACCGCCCCCGACGCCACGAGTGTACCTCGCCCCGAGACACAGCCCGGCCCC 516

Search completed: January 16, 2005, 09:30:30
Job time: 183.4 secs

This Page Blank (uspto)


```
/ NAME/KEY: unsure
/ LOCATION: 7155
/ OTHER INFORMATION: unknown
/ NAME/KEY: unsure
/ LOCATION: 7184
/ OTHER INFORMATION: unknown
/ NAME/KEY: unsure
/ LOCATION: 7957
/ OTHER INFORMATION: unknown
/ NAME/KEY: intron
/ LOCATION: (2997)...(3498)
/ OTHER INFORMATION: intron 1
/ NAME/KEY: intron:exon junction
/ LOCATION: (3498)...(3499)
/ OTHER INFORMATION: intron 1:exon 2
/ NAME/KEY: exon
/ LOCATION: (3508)...(3599)
/ OTHER INFORMATION: exon 2d
/ NAME/KEY: exon:intron junction
/ LOCATION: (3982)...(3983)
/ OTHER INFORMATION: exon 2a:intron 2a
/ NAME/KEY: intron:exon junction
/ LOCATION: (4205)...(4206)
/ OTHER INFORMATION: intron 2c:exon 3c
/ NAME/KEY: intron:exon junction
/ LOCATION: (4259)...(4260)
/ OTHER INFORMATION: intron 2d:exon 3d
/ NAME/KEY: exon
/ LOCATION: (4260)...(4328)
/ OTHER INFORMATION: exon 3d
/ NAME/KEY: intron:exon junction
/ LOCATION: (4632)...(4633)
/ OTHER INFORMATION: intron 3:exon 4
/ NAME/KEY: exon
/ LOCATION: (4914)...(5035)
/ OTHER INFORMATION: exon 5
/ NAME/KEY: intron
/ LOCATION: (5266)...(6293)
/ OTHER INFORMATION: intron 6
US-10-029-517-19
```

```
Query Match          63.7%; Score 38.2; DB 4; Length 8186;
Best Local Similarity 78.0%; Pred. No. 0.0014;
Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
```

```
QY 1 GGGTCGACTGCCCCCTCCGGCGCATGTGTGACCTCCTGACACAGCCAGCCCC 59
DB 3825 GGCTCCACCGCCCCCGCCAGCCCGGTGTCACTCGGCCCGGACACCGCCGCC 3883
```

RESULT 3

```
US-10-029-517-18
/ Sequence 18, Application US/10029517
/ Patent No. 6716627
/ GENERAL INFORMATION:
/ APPLICANT: Kenneth W. Dobie
/ TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
/ FILE REFERENCE: RTS-0352
/ CURRENT APPLICATION NUMBER: US/10/029,517
/ NUMBER OF SEQ ID NOS: 107
/ SEQ ID NO 18
/ LENGTH: 572
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (67)...(572)
US-10-029-517-18
```

```
Query Match          61.0%; Score 36.6; DB 4; Length 572;
Best Local Similarity 76.3%; Pred. No. 0.0033;
```

```
Matches 45; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
```

```
QY 1 GGGTCGACTGCCCCCTCCGGCGCATGTGTGACCTCCTGACACAGCCAGCCCC 59
DB 478 GGCTCCACCGCCCCCGCCAGCCCGGTGTCACTCGGCCCGGACACCGCCGCC 536
```

RESULT 4

```
US-10-029-517-3
/ Sequence 3, Application US/10029517
/ Patent No. 6716627
/ GENERAL INFORMATION:
/ APPLICANT: Kenneth W. Dobie
/ TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
/ FILE REFERENCE: RTS-0352
/ CURRENT APPLICATION NUMBER: US/10/029,517
/ NUMBER OF SEQ ID NOS: 107
/ SEQ ID NO 3
/ LENGTH: 1721
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (58)...(1605)
US-10-029-517-3
```

```
Query Match          61.0%; Score 36.6; DB 4; Length 1721;
Best Local Similarity 76.3%; Pred. No. 0.0038;
Matches 45; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
```

```
QY 1 GGGTCGACTGCCCCCTCCGGCGCATGTGTGACCTCCTGACACAGCCAGCCCC 59
DB 442 GGCTCCACCGCCCCCGCCAGCCCGGTGTCACTCGGCCCGGACACCGCCGCC 500
```

RESULT 5

```
US-10-029-517-101
/ Sequence 101, Application US/10029517
/ Patent No. 6716627
/ GENERAL INFORMATION:
/ APPLICANT: Kenneth W. Dobie
/ TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
/ FILE REFERENCE: RTS-0352
/ CURRENT APPLICATION NUMBER: US/10/029,517
/ NUMBER OF SEQ ID NOS: 107
/ SEQ ID NO 101
/ LENGTH: 518
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (67)...(518)
US-10-029-517-101
```

```
Query Match          60.3%; Score 36.2; DB 4; Length 518;
Best Local Similarity 77.2%; Pred. No. 0.0044;
Matches 44; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
```

```
QY 1 GGGTCGACTGCCCCCTCCGGCGCATGTGTGACCTCCTGACACAGCCAGCCCC 57
DB 462 GGCTCCACCGCCCCCGCCAGCCCGGTGTCACTCGGCCCGGACACCGCCGCC 518
```

RESULT 6

```
US-10-029-517-16
/ Sequence 16, Application US/10029517
/ Patent No. 6716627
/ GENERAL INFORMATION:
/ APPLICANT: Kenneth W. Dobie
/ TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
```

FILE REFERENCE: RTS-0352
CURRENT APPLICATION NUMBER: US/10/029,517
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 107
SEQ ID NO 16
LENGTH: 981
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: exon:exon junction
LOCATION: (464)...(465)
OTHER INFORMATION: exon 3b:exon 4
US-10-029-517-16

Query Match 60.3%; Score 36.2; DB 4; Length 981;
Best Local Similarity 77.2%; Pred. No. 0.0047;
Matches 44; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 1 GGGTCGACTGCCCCCTCGGCGCATGCTGACCTCCTGACACAGCCAGCC 57
Db 21 GGCTCACCAGCCCCCGCCAGCCATGTGTCTCCTCGGCGGACAGAGCCGCC 77

RESULT 7
US-10-029-517-102
Sequence 102, Application US/10029517
Patent No. 6716637

GENERAL INFORMATION:
APPLICANT: Kenneth W. Dobie
APPLICANT: Susan J. Myers
TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
FILE REFERENCE: RTS-0352
CURRENT APPLICATION NUMBER: US/10/029,517
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 107
SEQ ID NO 102
LENGTH: 3343
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
US-10-029-517-102

Query Match 60.3%; Score 36.2; DB 4; Length 3343;
Best Local Similarity 77.2%; Pred. No. 0.0055;
Matches 44; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 1 GGGTCGACTGCCCCCTCGGCGCATGCTGACCTCCTGACACAGCCAGCC 57
Db 1728 GGCTCACCAGCCCCCGCCAGCCATGTGTCTCCTCGGCGGACAGAGCCGCC 1784

RESULT 8
US-08-479-537A-1
Sequence 1, Application US/08479537A
Patent No. 5861381

GENERAL INFORMATION:
APPLICANT: CHAMBER, Pierre
APPLICANT: KIENY, Marie-Paule
APPLICANT: LATHY, Richard
APPLICANT: HAREVIENT, Maira
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
TREATMENT OR PREVENTION OF A MALIGNANT TUMOR
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22113-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/479,537A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 90/13101
FILING DATE: 23-OCT-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/FR91/00835
FILING DATE: 23-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/039,320
FILING DATE: 04-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/403,576
FILING DATE: 14-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Teakin, Robin L.
REGISTRATION NUMBER: 35,030
REFERENCE/DOCKET NUMBER: 017753-025
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 6192 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 58..120
FEATURE:
NAME/KEY: repeat_region
LOCATION: 439..5239
OTHER INFORMATION: /note= "The nucleotides spanning 439-5239 constitute a repeated region wherein the repeat is (OTHER INFORMATION: nucleotides and encodes 20 amino acids, 17 of which are fixed OTHER INFORMATION: The number of such repeats varies from 1 to 80."
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 121..6166
FEATURE:
NAME/KEY: repeat_region
LOCATION: 457
OTHER INFORMATION: /note= "Nucleotide 457 is X1 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCG
OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 487
OTHER INFORMATION: /note= "Nucleotide 487 is Y = NNN
OTHER INFORMATION: which is the codon for Thr or Asn wherein Thr = ACT, ACC, ACF
OTHER INFORMATION: or ACG; and Asn = AAT or AAC."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 496
OTHER INFORMATION: /note= "Nucleotide 496 is X2 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCG
OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
US-08-479-537A-1

Query Match 58.3%; Score 35; DB 2; Length 6192;
Best Local Similarity 74.6%; Pred. No. 0.015;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 1 GGGTCGACTGCCCCCTCGGCGCATGCTGACCTCCTGACACAGCCAGCC 59
Db 382 GGCTCACCAGCCCCCGCCAGCCATGTGTCTCCTCGGCGGACAGAGCCGCC 440

RESULT 9
US-09-083-116-1
; Sequence 1, Application US/09083116
; Patent No. 6203795
; GENERAL INFORMATION:
; APPLICANT: CHAMON, Pierre
; APPLICANT: KIENY, Marie-Paule
; APPLICANT: LATHÉ, Richard
; APPLICANT: HAREUVENI, Mara
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
; TREATMENT OR PREVENTION OF A MALIGNANT TUMOR
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/083,116
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/479,537
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/FR91/00835
; FILING DATE: 23-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/039,320
; FILING DATE: 04-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/403,576
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Teskin, Robin L.
; REGISTRATION NUMBER: 35,030
; REFERENCE/DOCKET NUMBER: 017753-025
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6192 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 58..120
; NAME/KEY: repeat_region
; LOCATION: 439..5239
; OTHER INFORMATION: /note="The nucleotides spanning
; 439-5239 constitute a repeated region wherein the repeat is 6
; OTHER INFORMATION: nucleotides and encodes 20 amino acids, 17 of which are fixed
; OTHER INFORMATION: The number of such repeats varies from 1 to 80."
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 121..6166
; NAME/KEY: repeat_region
; LOCATION: 457
; OTHER INFORMATION: /note="Nucleotide 457 is XI = NNN
; OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA

OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 487
; OTHER INFORMATION: /note="Nucleotide 487 is Y = NNN
; OTHER INFORMATION: which is the codon for Thr or Asn wherein Thr = ACT, ACC, ACA
; OTHER INFORMATION: or ACG; and Asn = AAT or AAC."
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 496
; OTHER INFORMATION: /note="Nucleotide 496 is X2 = NNN
; OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
; OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
US-09-083-116-1
Query Match 58.3%; Score 35; DB 3; Length 6192;
Best Local Similarity 74.6%; Pred. No. 0.015;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 1 GGGTGAAGTGGCCCTCCGCGCATGTGTGACCTCACTCTCTGACACAGAGCCGCCCC 59
DB 382 GGCTCCACACCCGCGCCAGCCGATGTCACTGACCCCGGAGCAACAGCCGCCCC 440
RESULT 10
US-09-134-916A-1
; Sequence 1, Application US/09134916A
; Patent No. 6328956
; GENERAL INFORMATION:
; APPLICANT: CHAMON, Pierre
; APPLICANT: KIENY, Marie-Paule
; APPLICANT: LATHÉ, Richard
; APPLICANT: HAREUVENI, Mara
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
; TREATMENT OR PREVENTION OF A MALIGNANT TUMOR
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/134,916A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/479,537
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: FR 90/13101
; FILING DATE: 23-OCT-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/FR91/00835
; FILING DATE: 23-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/039,320
; FILING DATE: 04-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/403,576
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Teskin, Robin L.
; REGISTRATION NUMBER: 35,030
; REFERENCE/DOCKET NUMBER: 017753-025
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 6192 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 58..120
FEATURE:
NAME/KEY: repeat_region
LOCATION: 439..5239
OTHER INFORMATION: /note= "The nucleotides spanning 439-5239 constitute a repeated region wherein the repeat is 6
OTHER INFORMATION: nucleotides and encodes 20 amino acids, 17 of which are fixed
OTHER INFORMATION: The number of such repeats varies from 1 to 80."
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 121..6166
FEATURE:
NAME/KEY: repeat_region
LOCATION: 457
OTHER INFORMATION: /note= "Nucleotide 457 is X1 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 487
OTHER INFORMATION: /note= "Nucleotide 487 is Y = NNN
OTHER INFORMATION: which is the codon for Thr or Asn wherein Thr = ACT, ACC, ACA
OTHER INFORMATION: or ACG; and Asn = AAT or AAC."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 496
OTHER INFORMATION: /note= "Nucleotide 496 is X2 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
US-09-134-916A-1
Query Match 58.3%; Score 35; DB 3; Length 6192;
Best Local Similarity 74.6%; Pred. No. 0.015;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 1 GGGTGACTGCCCTCCGGCGCATGTGTGACTGACTCTCTCTGACACAGCCAGCCCC 59
DB 382 GGCTCCACCACCCCGCCGCGCCGACGATGTCACTCAGCCCCGAGACAGCCAGCCCC 440
RESULT 11
US-08-479-537A-4
Sequence 4, Application US/08479537A
Patent No. 5861381
GENERAL INFORMATION:
APPLICANT: CHAMON, Pierre
APPLICANT: KIENY, Marie-Paule
APPLICANT: LATHE, Richard
APPLICANT: HAREUVEN, Mera
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
TREATMENT OR PREVENTION OF A MALIGNANT TUMOR
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/479,537A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 90/13101
FILING DATE: 23-OCT-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/FR91/00835
FILING DATE: 23-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/039,320
FILING DATE: 04-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/403,576
FILING DATE: 14-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Teskin, Robin L.
REGISTRATION NUMBER: 35,030
REFERENCE/DOCKET NUMBER: 017753-025
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-6620
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 6449 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 58..120
FEATURE:
NAME/KEY: repeat_region
LOCATION: 439..5239
OTHER INFORMATION: /note= "The nucleotides spanning 439-5239 constitute a repeated region wherein the repeat is 6
OTHER INFORMATION: nucleotides and encodes 20 amino acids, 17 of which are fixed
OTHER INFORMATION: The number of such repeats varies from 1 to 80."
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 121..5661
FEATURE:
NAME/KEY: repeat_region
LOCATION: 457
OTHER INFORMATION: /note= "Nucleotide 457 is X1 = NNN
OTHER INFORMATION: which is the codon for Thr or Ala wherein Pro = CCT, CCC, CCA
OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 487
OTHER INFORMATION: /note= "Nucleotide 487 is Y = NNN
OTHER INFORMATION: which is the codon for Thr or Asn wherein Thr = ACT, ACC, ACA
OTHER INFORMATION: or ACG; and Asn = AAT or AAC."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 496
OTHER INFORMATION: /note= "Nucleotide 496 is X2 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
US-08-479-537A-4
Query Match 58.3%; Score 35; DB 2; Length 6449;
Best Local Similarity 74.6%; Pred. No. 0.015;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 1 GGGTGACTGCCCTCCGGCGCATGTGTGACTGACTCTCTCTGACACAGCCAGCCCC 59
DB 382 GGCTCCACCACCCCGCCGCGCCGACGATGTCACTCAGCCCCGAGACAGCCAGCCCC 440
RESULT 12

US-09-083-116-4
Sequence 4, Application US/09083116
Patent No. 6203795
GENERAL INFORMATION:
APPLICANT: CHAMBER, Pierre
APPLICANT: KIENY, Marie-Paule
APPLICANT: LATHE, Richard
APPLICANT: HAREVENT, Mara
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
TREATMENT OR PREVENTION OF A MALIGNANT TUMOR
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22113-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/083,116
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/479,537
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/FR91/00835
FILING DATE: 23-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/039,320
FILING DATE: 04-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/403,576
FILING DATE: 14-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Teskin, Robin L.
REGISTRATION NUMBER: 35,030
REFERENCE/DOCKET NUMBER: 017753-025
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 6449 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 58..120
FEATURE:
NAME/KEY: repeat_region
LOCATION: 439..5239
OTHER INFORMATION: /note= "The nucleotides spanning
OTHER INFORMATION: 439-5239 constitute a repeated region wherein the repeat is
OTHER INFORMATION: nucleotides and encodes 20 amino acids, 17 of which are fixed
OTHER INFORMATION: The number of such repeats varies from 1 to 80."
NAME/KEY: mat_peptide
LOCATION: 121..5661
FEATURE:
NAME/KEY: repeat_region
LOCATION: 457
OTHER INFORMATION: /note= "Nucleotide 457 is X1 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
FEATURE:

NAME/KEY: repeat_region
LOCATION: 487
OTHER INFORMATION: /note= "Nucleotide 487 is Y = NNN
OTHER INFORMATION: which is the codon for Thr or Asn wherein Thr = ACT, ACC, ACA
OTHER INFORMATION: or ACG; and Asn = AAT or AAC."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 496
OTHER INFORMATION: /note= "Nucleotide 496 is X2 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
US-09-083-116-4
Query Match
Best Local Similarity 58.3%; Score 35; DB 3; Length 6449;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 1 GGGTGACTGCCCTCCGCGCATGTGTGACCTCAGCTCCTGACACAGGCCAGCCCC 59
DB 382 GGCTCCACACCCCGCGCATGTGTGACCTCAGCTCCTGACACAGGCCAGCCCC 440
RESULT 13
US-09-134-916A-4
Sequence 4, Application US/09134916A
Patent No. 6328956
GENERAL INFORMATION:
APPLICANT: CHAMBER, Pierre
APPLICANT: KIENY, Marie-Paule
APPLICANT: LATHE, Richard
APPLICANT: HAREVENT, Mara
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
TREATMENT OR PREVENTION OF A MALIGNANT TUMOR
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22113-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/134,916A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/479,537
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: FR 90/13101
FILING DATE: 23-OCT-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/FR91/00835
FILING DATE: 23-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/039,320
FILING DATE: 04-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/403,576
FILING DATE: 14-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Teskin, Robin L.
REGISTRATION NUMBER: 35,030
REFERENCE/DOCKET NUMBER: 017753-025
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:


```

; LENGTH: 6449 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: s1g_peptide
; LOCATION: 58..120
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 439..5239
; OTHER INFORMATION: /note= "The nucleotides spanning
; OTHER INFORMATION: 439-5239 constitute a repeated region wherein the repeat is 6
; OTHER INFORMATION: nucleotides and encodes 20 amino acids, 17 of which are fixed
; OTHER INFORMATION: The number of such repeats varies from 1 to 80."
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 121..5661
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 457
; OTHER INFORMATION: /note= "Nucleotide 457 is X1 = NNN
; OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
; OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 487
; OTHER INFORMATION: /note= "Nucleotide 487 is Y = NNN
; OTHER INFORMATION: which is the codon for Thr or Asn wherein Thr = ACT, ACC, ACA
; OTHER INFORMATION: or ACG; and Asn = AAT or AAC."
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 496
; OTHER INFORMATION: /note= "Nucleotide 496 is X2 = NNN
; OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
; OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
; US-09-134-916A-4

Query Match      58.3%; Score 35; DB 3; Length 6449;
Best Local Similarity 74.6%; Pred. No. 0.015;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 1 GGGTGACTGCCCCCTCCGGCGCATGTGTGACCTGAGCTTGACACAGGCGCC 59
Db 382 GGCTCCACGCCCCCGCGCAGCCACGATGTCACTTCAAGCCCGACACAGCCAGCCCC 440

RESULT 14
US-09-475-947A-246
; Sequence 246, Application US/09475947A
; Patent No. 6472154
; GENERAL INFORMATION:
; APPLICANT: Garner, Harold R.
; APPLICANT: Wren, Jonathan D.
; APPLICANT: Minna, John D.
; TITLE OF INVENTION: Polymorphic Repeats in Human Genes
; FILE REFERENCE: UTS0067
; CURRENT APPLICATION NUMBER: US/09/475,947A
; CURRENT FILING DATE: 1999-12-31
; NUMBER OF SEQ ID NOS: 346
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 246
; LENGTH: 60
; TYPE: DNA
; ORGANISM: human
; US-09-475-947A-246

Query Match      56.3%; Score 33.8; DB 4; Length 60;
Best Local Similarity 77.4%; Pred. No. 0.02;
Matches 41; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 1 GGGTGACTGCCCCCTCCGGCGCATGTGTGACCTGAGCTTGACACAGGCGCC 53

```

```

Db 7 GGCTCCACGCCCCCGCGCAGCCACGATGTCACTTCAAGCCCGACACAGGCGCC 59

RESULT 15
US-09-722-139-1/c
; Sequence 1, Application US/09722139
; Patent No. 6355471
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Freedman, Richard
; TITLE OF INVENTION: No. 6355471e1 motor proteins and methods for
; FILE REFERENCE: their use
; FILE REFERENCE: 1055
; CURRENT APPLICATION NUMBER: US/09/722,139
; CURRENT FILING DATE: 2000-11-24
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 4176
; TYPE: DNA
; ORGANISM: Human
; US-09-722-139-1

Query Match      41.3%; Score 24.8; DB 3; Length 4176;
Best Local Similarity 63.3%; Pred. No. 29;
Matches 38; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy 1 GGGTGACTGCCCCCTCCGGCGCATGTGTGACCTGAGCTTGACACAGGCGCCCA 60
Db 3073 GGGTGAGTGCTCTGTGACGCGCAGATGTCTCTCCAGCCTGCGCGCGCTCCA 3014

```

Search completed: January 16, 2005, 03:01:44
Job time : 36.2 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using bw model

Run on: January 15, 2005, 19:43:55 : Search time 1550.6 Seconds
(without alignments)
1410.024 Million cell updates/sec

Title: US-10-057-136-10

Perfect score: 60

Sequence: 1 gggctgcagcgcgcctccgcgc.....ctgacacaaagccagcccca 60

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39.8	66.3	330	1	A1925867
2	39.8	66.3	1262	5	BQ935496
3	39.8	66.3	1349	5	BQ935496
4	38.2	63.7	604	4	BM791359
5	38.2	63.7	877	4	BM791359
6	38.2	63.7	959	6	CA489836
7	38.2	63.7	1113	5	BU148487
8	38.2	63.7	1130	5	BU542996
9	38.2	63.7	1234	5	BQ936898
10	38.2	63.7	1268	5	BQ936898
11	38.2	63.7	1343	5	BQ936898
12	38.2	63.7	1430	5	BQ936898
13	38.2	63.7	1531	5	BQ936898
14	36.2	60.3	1536	5	BQ936898
15	36.2	60.3	1536	5	BQ936898
16	35.2	58.7	166	7	T27692
17	35.2	58.7	166	7	T27692
18	35.2	58.7	166	7	T27692
19	35.2	58.7	166	7	T27692
20	35.2	58.7	166	7	T27692
21	30.2	50.3	1678	4	BG775565
22	29.8	49.3	619	4	BI260921
23	29.6	49.3	619	4	BI260921
24	29.2	48.7	548	2	BE706360

25	27.8	46.3	528	4	BM130288
26	27.8	46.3	535	4	BM130288
27	27.8	46.3	538	4	BM130288
28	27.2	45.3	270	5	BY148470
29	27.2	45.3	270	5	BY148470
30	26.8	44.7	300	8	AO310138
31	26.8	44.7	300	8	AO310138
32	26.8	44.7	300	8	AO310138
33	26.8	44.7	300	8	AO310138
34	26.8	44.7	300	8	AO310138
35	26.4	44.0	617	7	CD218221
36	26.4	44.0	617	7	CD218221
37	26.4	44.0	838	9	CG306143
38	26.4	44.0	847	3	AY109264
39	26.4	44.0	898	9	CG053439
40	26.4	44.0	937	7	CK370897
41	26.4	44.0	982	9	CG249534
42	26.2	43.7	1172	7	CK209164
43	26.2	43.7	1298	6	CA803305
44	26.2	43.7	630	5	BO199203
45	26.2	43.3	358	5	BY466019

ALIGNMENTS

RESULT 1
LOCUS A1925867/c
DEFINITION WO2004.01 NCI CGAP Paul Homo sapiens CDNA clone IMAGE:2455879.3
similar to contains element MSRI MSRI repetitive element ;, mRNA
sequence.
ACCESSION A1925867
VERSION A1925867.1
KEYWORDS GI:5661831
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS
TITLE
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Life Technologies catalog #: 11548-013
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/BLN at:
www.bio.lnl.gov/birdp/image/image.html
Insert length: 578 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 305.
Location/Qualifiers
1..330
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2455879"
/issue_type="adenocarcinoma"
/lab_host="DH10B"
/clone_lib="NCI-CGAP Paul"
/note="Organ: pancreas; Vector: pCMV-Sport6; Site 1: SalI; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.72 kb. Life Technologies catalog #: 11548-013"

JOURNAL

COMMENT

FEATURES

source

ORIGIN

Query Match 66.3%; Score 39.8; DB 1; Length 330;
Best Local Similarity 79.7%; Pred. No. 0.028;
Matches 47; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

OY	1	GGGTCGACTGCCTCCCGCAGTAGTGTCGACCTTCACTCTGTACACACAAGGCCAGCCCC	59
Dp	256	GGCTCCACCGCCCCCCCAGCCACAGGTGTCACTTCGCTCCGAGCACACGAGCCGCGCCCC	198
RESULT 2			
LOCUS	BQ935496		
DEFINITION	BQ935496	1262 bp	mRNA
ACCSSION	AEENOCOURJ_8776475 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6382088		EST-21-AUG-2002
VERSION	BQ935496		
KEYWORDS	BQ935496.1 GI:22350879		
SOURCE	EST.		
ORGANISM	Homo sapiens (human)		
	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1 (bases 1 to 1262)		
AUTHORS	NIH-MGC http://mgi.nci.nih.gov/.		
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: Robert Strausberg, Ph.D.		

FEATURES

Source

ORIGIN

Query Match	66.3%	Score 39.8;	DB 5;	Length 1262;
Best Local Similarity	79.7%;	Pred. No. 0.029;		
Matches 47;	Conservative 0;	Mismatches 12.		

[illegible]

RESULT 3	BU152566	1349 bp	mRNA	linear	EST 03-SEP-2002
LOCUS	BU152566				
DEFINITION	BU152566				
	AGENE	OUT 8670683	NIH_MGC_40	Homo sapiens	cdna clone IMAGE:6380559
		5', mRNA	sequence.		
ACCESSION	BU152566				
VERSION	BU152566.1	GI:2266098			
KEYWORDS	EST.				
SOURCE	Homo sapiens				
ORGANISM	Homo sapiens (human)				

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 1349)				
		NIH-MGC http://mgc.ncbi.nlm.nih.gov/		
		National Institutes of Health, Mammalian Gene Collection (MGC)		
		Unpublished (1999)		
		Contact: Robert Strauberg, Ph.D.		

ORIGIN

Query Match	66.3%	Score 39.8;	DB 5;	Length 1349;
Best Local Similarity	79.7%	Pred. No. 0.029;		
Matches 47;	Conservative 0;	Mismatches 12;	Indels 0;	Gaps 0;

[illegible]

RESULT	4
BM791359	
LOCUS	604 bp mRNA linear EST 05-MAR-2002
DEFINITION	K-EST0071342 S21SNU520 Homo sapiens cDNA clone S21SNU520-14-A06 5'
ACCESSION	BM791359
VERSION	BM791359.1 GI:19139591
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens

REFERENCE AUTHORS	TITLE	JOURNAL	COMMENT
Mammalia: Eutheria; Primates; Carnivora; Eutelestomi 1 (bases 1 to 604)	Oh, K.Y., Hahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Kim, M.F.	21C Frontier Research EST Project 2001	
Kim, Y.S., Sohn, H.Y., Kim, J.M., Park, H.S., Kim, S. and Kim, Y.S.	21C Frontier Research EST Project 2001	Unpublished (2002)	Contact: Kim YS

FEATURES
source

RESULT 7
LOCUS BU148487
DEFINITION BU148487 1113 bp mRNA linear EST 03-SEP-2002
AGENCOURT 8670479 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6380642
5', mRNA sequence.
ACCESSION BU148487
VERSION BU148487.1 GI:22662019
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1113)
NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LNCM2569 row: j column: 03
High quality sequence stop: 235.
Location/Qualifiers
1..1113
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6380642"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
/note="Organ: prostate; Vector: pOT7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCAAGG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC library."

Query Match 63.7%; Score 38.2; DB 5; Length 1113;
Best Local Similarity 78.0%; Pred. No. 0.092;
Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
QY 1 GGGTCGACCTCCCTCCGCGCATGTGACCTCCTCGACACAGCCGCCCC 59
DB 30 GGCTCCACCGCCCCCGACCGCATGTGACCTCCTCGACACAGCCGCCCC 88

RESULT 8
LOCUS BU542996
DEFINITION BU542996 1130 bp mRNA linear EST 13-SEP-2002
AGENCOURT 10338707 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6574979
5', mRNA sequence.
ACCESSION BU542996
VERSION BU542996.1 GI:22853479
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1130)
NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LNCM2771 row: c column: 11
High quality sequence start: 27
High quality sequence stop: 246.
Location/Qualifiers
1..1130
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6574979"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
/note="Organ: prostate; Vector: pOT7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCAAGG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC library."

ORIGIN
Query Match 63.7%; Score 38.2; DB 5; Length 1130;
Best Local Similarity 78.0%; Pred. No. 0.092;
Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
QY 1 GGGTCGACCTCCCTCCGCGCATGTGACCTCCTCGACACAGCCGCCCC 59
DB 69 GGCTCCACCGCCCCCGACCGCATGTGACCTCCTCGACACAGCCGCCCC 127

RESULT 9
LOCUS B0936898
DEFINITION B0936898 1234 bp mRNA linear EST 21-AUG-2002
AGENCOURT 8919110 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6383855
5', mRNA sequence.
ACCESSION B0936898
VERSION B0936898.1 GI:22352281
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1234)
NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LNCM2577 row: c column: 24
High quality sequence stop: 245.
Location/Qualifiers
1..1234
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6383855"

FEATURES
source

ORIGIN

Qy 1 GGGTCGACTGCCCCCTCCGGCGCATGATGTGACTCTAGCTCCTTACACAGGCCCAAGCCCC 53

Db 30 GGCTCCACGCGCCCCCGGAGCCAGCGGTGCACTTGTGCCCCGGGACACAGAGCGGCCCC 88

VERSION	BQ943554.1	GI:22359032
KEYWORDS	EST.	
SOURCE	Homo sapiens	(human)
ORGANISM	Homo sapiens	

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 1268)	NIH-MGC	http://mgc.nci.nih.gov/	National Institutes of Health, Mammalian Gene Collection (MGC)	Unpublished (1999)
	Contact:	Robert Strausberg, Ph.D.		

cdna library Preparation: Rubin Laboratory
cdna library Arrayed by: The I.M.A.G.E. Consortium (LINTL)
DNA sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LINTL at:
<http://image.llnl.gov>
Plate: LINC5579 row: b column: 21
High quality sequence: scdp: 177.

FEATURES	Location/Qualifiers
source	1. .1268

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon: 9606"
/clone="IMAGE:6384308"
/risuse_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone.lib="NIH MGC 40"
/notes="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGGCACGGG(6). Library constructed by
ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life
Technologies).
Note: This is a NIH MGC library."

```

ORIGIN

Query Match	63.7%	Score 38.2;	DB 5;	Length 1268;
Best Local Similarity	78.0%	Pred. No. 0.093;		
Matches 46;	Conservative 0;	Mismatches 13;	Indels 0;	Gaps 0;

QY 1 GGGTCGACTGCCCTCCGGCGCATGTTGTGACCTCAGCTCTGACACAAAGGCCAGCCC 59

Db 90 GGCTCCACCGCCCCCAGCCACGGTGTACCTCG6CCCCGGAACAGGCCGGCCCC 148

RESULT	LOCUS	DEFINITION
11 BQ920055	BQ920055	1343 bp mRNA linear EST 20-AUG-2002 ACENSCOURT_10031674 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6481605 5' mRNA sequence.

VERSION	BQ920055.1	GI:22334753
KEYWORDS	EST.	

ORGANISM

REFERENCE	1 (pages 1 to 1343)
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC).
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D. rstraus@nih.gov

CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINT)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LINT at:
<http://image.llnl.gov>
Plate: LINC2663 row: e column: 02
High quality sequence start: 56
High quality sequence stop: 237.

FEATURES

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="takon:9606"
/clone="IMAGE:6481609"
/issue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_1lb="NIH_MGC_40"
/notes="Organ: prostate; Vector: pOT87; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAGAC(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

```

ORIGIN

Query Match	63.7%	Score 38.2;	DB 5;	Length 1343;
Best Local Similarity	78.0%;	Pred. No. 0.093;		
Matches 46;	Conservative 0;	Mismatches 13;	Indels 0;	Gaps 0;

Qy 1 GGGTGAATGCCCCCTCCGGCGCATGGTGTGACCTCAGCTCTTGACACAAAGCCAGCCCC 59
Db 72 GGCTTCACCGCCCCCCCAAGCAGGTGTACCTCTGCCCCGGAACACAGAGCCGCCCC 13

RESULT 12					
LOCUS	BU542790	1420 bp	mRNA	linear	EST 13-SEP-2003
DEFINITION	ABENCCOURT 10334841 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:65747725				
DESCRIPTION	5', mRNA sequence.				

ACCESSION	BU542790
VERSION	BU542790.1
	GI:22853273

KEYWORDS

SOURCE ORGANISM	
Homo sapiens (human)	
Homo sapiens	

REFERENCE 1 (bases 1 to 1420)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-r@mail.nih.gov
 Tissue Procurement: DCTD/DRP

CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: LNCM2770 row: h column: 21
 High quality sequence stop: 288.

FEATURES
SOURCE
 1..1420
 Location/Qualifiers

/organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6574725"
 /tissue_type="carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_40"
 /note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
 Site_2: EcoRI; CDNA made by oligo-dT priming;
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCACGAG(G). Library constructed by
 Ling Hong in the laboratory of Gerald M. Rubin (University
 of California, Berkeley) using ZAP-CDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)."
 Note: this is a NIH_MGC Library."

ORIGIN

Query Match
 Best Local Similarity 63.7%; Score 38.2; DB 5; Length 1420;
 Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 1 GGCTGACTGCCCCCTCCGCGCATGTGTGACCTCCTCTGACACAGCCAGCCCC 59
 19 GGCTCACCGCCCCCTCCGCGCATGTGTGACCTCCTCTGACACAGCCAGCCCC 77

RESULT 13
LOCUS BUS43309 1531 bp mRNA linear EST 13-SEP-2002
DEFINITION AGENCOURT_10327072 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6575349
ACCESSION BUS43309
VERSION BUS43309.1 GI:22853792
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 1 (bases 1 to 1531)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-r@mail.nih.gov
 Tissue Procurement: DCTD/DRP

CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: LNCM2772 row: b column: 21
 High quality sequence stop: 166.

FEATURES
SOURCE
 1..1531
 Location/Qualifiers

/organism="Homo sapiens"
 /mol_type="mRNA"

ORIGIN

Query Match
 Best Local Similarity 63.7%; Score 38.2; DB 5; Length 1511;
 Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 1 GGCTGACTGCCCCCTCCGCGCATGTGTGACCTCCTCTGACACAGCCAGCCCC 59
 30 GGCTCACCGCCCCCTCCGCGCATGTGTGACCTCCTCTGACACAGCCAGCCCC 88

RESULT 14
LOCUS BG774910 981 bp mRNA linear EST 15-MAY-2001
DEFINITION 602649832P1 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:4761054 5',
ACCESSION BG774910
VERSION BG774910.1 GI:14045227
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 1 (bases 1 to 981)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-r@mail.nih.gov
 Tissue Procurement: DCTD/DRP

CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: LNCM1612 row: o column: 07
 High quality sequence stop: 874.

FEATURES
SOURCE
 1..981
 Location/Qualifiers

/organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4761054"
 /tissue_type="carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_40"
 /note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
 Site_2: EcoRI; CDNA made by oligo-dT priming;
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCACGAG(G). Library constructed by
 Ling Hong in the laboratory of Gerald M. Rubin (University
 of California, Berkeley) using ZAP-CDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)."
 Note: this is a NIH_MGC Library."

ORIGIN

Query Match
 Best Local Similarity 60.3%; Score 36.2; DB 4; Length 981;
 Matches 44; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

OY 1 GGGTCGACTGCCCCCTCGGCGCATGGTGTGACCTGACCTGACACACAGCCAGCC 57
 DB 21 GGCTTCACCGCCCCCGCCAGCCATGGTGTGATCCTCGGCCCGGACACAGAGCCCGCC 77

RESULT 15
 BO923149 1536 bp mRNA linear EST 20-AUG-2002
 LOCUS BO923149 5', mRNA sequence.
 DEFINITION AGENCOURT_8929207 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6484568
 ACCESSION BO923149
 VERSION BO923149.1 GI:22338180
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 1536)
 NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: DCTD/DTF

CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.lnl.gov
 Plate: LCM2670 row: P column: 09
 High quality sequence stop: 287.

FEATURES
 source Location/Qualifiers

1..1536
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6484568"
 /tissue_type="carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 40"
 /note="Organ: prostate; Vector: pORF7; Site 1: XhoI;
 Site 2: EcoRI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCAAGAG(G). Library constructed by
 Ling Hong in the laboratory of Gerald M. Rubin (University
 of California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC library."

ORIGIN

Query Match 60.3%; Score 36.2; DB 5; Length 1536;
 Best Local Similarity 77.2%; Pred. No. 0.4;
 Matches 44; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
 OY 1 GGGTCGACTGCCCCCTCGGCGCATGGTGTGACCTGACCTGACACACAGCCAGCC 57
 DB 18 GGCTTCACCGCCCCCGCCAGCCATGGTGTGATCCTCGGCCCGGACACAGAGCCCGCC 74

Search completed: January 16, 2005, 02:55:38
 Job time : 1552.6 secs

This Page Blank (uspto)

XX Sequence 60 BP; 10 A; 24 C; 17 G; 9 T; 0 U; 0 Other;
SQ

Query Match 100.0%; Score 60; DB 2; Length 60;
Best Local Similarity 100.0%; Pred. No. 1.8e-10;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGTCGACAGTCCCTCCGGCGCATGTGTGACCTCAGCTCTGACACAGGCCAGCCCCA 60
DB 1 GGCTCCAGCTGCCCTCCGGCGCATGTGTGACCTCAGCTCTGACACAGGCCAGCCCCA 60

RESULT 2
AD157666
ID AD157666 standard; cDNA; 4144 BP.

XX AD157666;

XX 22-APR-2004 (first entry)

DE Human breast specific nucleic acid (BSNA) #37.

XX Human; breast specific nucleic acid; BSNA; gene; ss; metastasis;
KW breast cancer; cytostatic.

OS Homo sapiens.

XX WO2003106648-A2.

XX 24-DEC-2003.

XX 16-JUN-2003; 2003WO-US018934.

XX 14-JUN-2002; 2002US-0389327P.

PA (DIAD-) DIADEXUS INC.

PI Salceda S, Macina RA, Turner LR, Sun Y, Liu C;

DR WPI; 2004-082185/08.
P-PSDB; AD157740.

PT Novel isolated polypeptide comprising breast specific protein sequences,
useful for diagnosing or monitoring presence and metastases of breast
cancer in patient.

PS Claim 1; SEQ ID NO 37; 370pp; English.

XX The invention relates to human breast specific nucleic acids (BSNA) and
the breast specific proteins (BSP) they encode. The nucleic acids are
useful for determining the presence of a BSNA in a sample which involves
contacting the sample with a BSNA under conditions in which the BSNA will
selectively hybridise to a BSNA in the sample, and detecting the
hybridisation. The nucleic acids are useful for determining the presence
of a BSP in a sample which involves contacting the sample with suitable
reagent under conditions in which the reagent will selectively interact
with the BSP, and detecting the interaction of the reagent with a BSP in
the sample. The nucleic acids and proteins are useful for diagnosing or
monitoring the presence and metastases of breast cancer in a patient,
which involves determining an amount of nucleic acid or protein and
comparing the determined amount of nucleic acid or protein in the
of the patient to the amount of a breast specific marker in a normal
control, where a difference in the determined amount in the sample
compared to the amount in the control is associated with the presence of
breast cancer. The sequences are useful for treating a patient with
breast cancer, involving administering a composition consisting of a BSNA
or a BSP to a patient, where the administration induces an immune
response against the breast cancer cell expressing the BSNA or BSP. This
sequence represents a human BSNA of the invention.

SQ Sequence 4144 BP; 634 A; 1896 C; 1058 G; 552 T; 0 U; 4 Other;

Query Match

65.7%; Score 39.4; DB 12; Length 4144;

Best Local Similarity 78.0%; Pred. No. 0.0022;
Matches 46; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

QY 1 GGGTCGACAGTCCCTCCGGCGCATGTGTGACCTCAGCTCTGACACAGGCCAGCCCC 59
DB 693 GGCTCCAGCTGCCCTCCGGCGCATGTGTGACCTCAGCTCTGACACAGGCCAGCCCC 751

RESULT 3

ID AA48316 standard; cDNA; 60 BP.

XX AA48316;

XX 20-NOV-1998 (first entry)

DE Nucleotide sequence encoding MUC1 tandem repeat unit.

XX ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
KW tumour; tumour-associated antigen.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 1..60
FT /tag= A
FT /product= "MUC1 tandem repeat unit"

XX WO9837095-A2.

XX 27-AUG-1998.

XX 24-FEB-1998; 98WO-US003693.

XX 24-FEB-1997; 97US-0038253P.

PA (THER-) THERION BIOLOGICS CORP.
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA (DAND) DANA FARNBER CANCER INST INC.

PI Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;

DR WPI; 1998-467492/40.
P-PSDB; AAW77229.

PT New recombinant pox virus for tumour therapy - comprises DNA encoding an
immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.

XX Example 1; Page 20; 42pp; English.

XX The MUC1 tandem repeat unit was used to create an immunogenic mini-MUC1
fragment for inclusion in a recombinant pox virus (RPV). The RPV was used
in a pharmaceutical composition also containing an immunomodulator to
generate MUC1 specific cytotoxic T-lymphocytes. The recombinant pox virus
therefore encodes an immunogenic MUC1 fragment that does not undergo
significant genetic deletion, thereby providing an unexpectedly stable
and immunogenic pox virus. They can be used to prevent or treat tumours
expressing MUC1 tumour-associated antigens

XX SQ Sequence 60 BP; 7 A; 33 C; 16 G; 4 T; 0 U; 0 Other;

Query Match 63.7%; Score 38.2; DB 2; Length 60;
Best Local Similarity 78.0%; Pred. No. 0.0032;
Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 GGGTCGACAGTCCCTCCGGCGCATGTGTGACCTCAGCTCTGACACAGGCCAGCCCC 59
DB 1 GGCTCCAGCTGCCCTCCGGCGCATGTGTGACCTCAGCTCTGACACAGGCCAGCCCC 59

RESULT 4

AA48316 standard; cDNA; 309 BP.

```
XX AC AAN90579;
XX XX
DT 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 04-DEC-1989 (first entry)
XX XX
DE PDF9.3 cDNA insert.
XX XX
KM PDF9.3; human DF3 breast carcinoma-associated antigen epitope.
XX OS Human MCF-7 breast carcinoma cells; 'ATCC HTB22'.
XX PN WO9907107-A.
XX PD 10-AUG-1989.
XX PF 29-JAN-1988; 88US-00149831.
XX PR 29-JAN-1988; 88US-00149831.
XX PA (DANA-) DANA-FARBER CANCER.
XX PI Kufe DW;
XX DR MPI: 1989-248989/34.
XX DR P-PSDB; AAP91045, AAP91053, AAP91054, AAP90146.
XX PT Recombinant polypeptide(s) - contains DF 3 breast carcinoma antigen
XX PT epitope and useful as assay reagents, and encoding DNA sequences.
XX PS Claim 1, Fig 4; 31pp; English.
XX CC The sequence encodes a carbohydrate-free polypeptide contg. DF3 breast
XX CC carcinoma antigen epitope. Useful as a competitive binding assay reagent
XX CC and improves diagnosis. The cDNA consists of nearly identical 60 BP
XX CC tandem repeats which are 85% GC-rich. See also AAP91045, AAP91046,
XX CC AAP91053, AAP91054. (Updated on 25-MAR-2003 to correct PA field.)
XX CC (Updated on 27-AUG-2003 to correct OS field.)
XX SQ Sequence 309 BP; 21 A; 92 C; 160 G; 36 T; 0 U; 0 Other;
XX XX
Query Match 63.7%; Score 38.2; DB 1; Length 309;
Best Local Similarity 78.0%; Pred. No. 0.0039;
Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
QY 1 GGGTCGACTGCCCTCCGGCGCATGTGTGACCTCAGCTCTGACACAGCCAGCCCC 59
DB 266 GGCTCCACCGCCGCCCCGACGCCGTCACCTCGCGCCCGGACACCGCGGCC 208
RESULT 5
ADIS7712
ID ADIS7712 standard; cDNA; 1194 BP.
XX AC ADIS7712;
XX DT 22-APR-2004 (first entry)
XX DE Human breast specific nucleic acid (BSNA) #83.
XX KM Human; breast specific nucleic acid; BSNA; gene; ss; metastasis;
XX KM breast cancer; cytostatic.
XX OS Homo sapiens.
XX PN WO2003106648-A2.
XX PD 24-DEC-2003.
XX PF 16-JUN-2003; 2003WO-US018934.
XX PR 14-JUN-2002; 2002US-0389327P.
XX PI 14-JUN-2002; 2002US-0389327P.
```

```
XX PA (DIAD-) DIADEXUS INC.
XX XX
XX PI Salceda S, Macina RA, Turner LR, Sun Y, Liu C;
XX DR MPI; 2004-082185/08.
XX DR P-PSDB; ADIS7782.
XX XX
PT Novel isolated polypeptide comprising breast specific protein sequences,
PT useful for diagnosing or monitoring presence and metastases of breast
PT cancer in patient.
XX PS Claim 1; SEQ ID NO 83; 370pp; English.
XX XX
CC The invention relates to human breast specific nucleic acids (BSNA) and
CC the breast specific proteins (BSP) they encode. The nucleic acids are
CC useful for determining the presence of a BSNA in a sample which involves
CC contacting the sample with a BSNA under conditions in which the BSNA will
CC selectively hybridise to a BSNA in the sample, and detecting the
CC hybridisation. The nucleic acids are useful for determining the presence
CC of a BSP in a sample which involves contacting the sample with suitable
CC reagent under conditions in which the reagent will selectively interact
CC with the BSP, and detecting the interaction of the reagent with a BSP in
CC the sample. The nucleic acids and proteins are useful for diagnosing or
CC monitoring the presence and metastases of breast cancer in a patient,
CC which involves determining an amount of nucleic acid or protein and
CC comparing the determined amount of nucleic acid or protein in the sample
CC of the patient to the amount of a breast specific marker in a normal
CC control, where a difference in the determined amount in the sample
CC compared to the amount in the control is associated with the presence of
CC breast cancer. The sequences are useful for treating a patient with
CC breast cancer, involving administering a composition consisting of a BSNA
CC or a BSP to a patient, where the administration induces an immune
CC response against the breast cancer cell expressing the BSNA or BSP. This
CC sequence represents a human BSNA of the invention.
XX SQ Sequence 1194 BP; 243 A; 438 C; 275 G; 238 T; 0 U; 0 Other;
XX XX
Query Match 63.7%; Score 38.2; DB 12; Length 1194;
Best Local Similarity 78.0%; Pred. No. 0.0047;
Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
QY 1 GGGTCGACTGCCCTCCGGCGCATGTGTGACCTCAGCTCTGACACAGCCAGCCCC 59
DB 596 GGCTCCACCGCCGCCCCGACGCCGTCACCTCGCGCCCGGACACCGCGGCC 654
RESULT 6
ADIS7693
ID ADIS7693 standard; cDNA; 1378 BP.
XX AC ADIS7693;
XX DT 22-APR-2004 (first entry)
XX DE Human breast specific nucleic acid (BSNA) #64.
XX KM Human; breast specific nucleic acid; BSNA; gene; ss; metastasis;
XX KM breast cancer; cytostatic.
XX OS Homo sapiens.
XX PN WO2003106648-A2.
XX PD 24-DEC-2003.
XX PF 16-JUN-2003; 2003WO-US018934.
XX PR 14-JUN-2002; 2002US-0389327P.
XX PA (DIAD-) DIADEXUS INC.
XX PI Salceda S, Macina RA, Turner LR, Sun Y, Liu C;
```

XX WPI: 2004-082185/08.
DR P-PSDB; AD157765.

XX Novel isolated polypeptide comprising breast specific protein sequences,
PT useful for diagnosing or monitoring presence and metastases of breast
PT cancer in patient.

XX Claim 1; SEQ ID NO 64; 370pp; English.

XX The invention relates to human breast specific nucleic acids (BSNA) and
XX the breast specific proteins (BSP) they encode. The nucleic acids are
XX useful for determining the presence of a BSNA in a sample which involves
XX contacting the sample with a BSNA under conditions in which the BSNA will
XX selectively hybridize to a BSNA in the sample, and detecting the
XX hybridization. The nucleic acids are useful for determining the
XX of a BSP in a sample which involves contacting the sample with suitable
XX reagent under conditions in which the reagent will selectively interact
XX with the BSP, and detecting the interaction of the reagent with a BSP in
XX the sample. The nucleic acids and proteins are useful for diagnosing or
XX monitoring the presence and metastases of breast cancer in a patient,
XX which involves determining an amount of nucleic acid or protein and
XX comparing the determined amount of nucleic acid or protein in the sample
XX of the patient to the amount of a breast specific marker in a normal
XX control, where a difference in the determined amount in the sample
XX compared to the amount in the control is associated with the presence of
XX breast cancer. The sequences are useful for treating a patient with
XX breast cancer, involving administering a composition consisting of a BSNA
XX or a BSP to a patient, where the administration induces an immune
XX response against the breast cancer cell expressing the BSNA or BSP. This
XX sequence represents a human BSNA of the invention.

XX Sequence 1378 BP; 295 A; 489 C; 316 G; 278 T; 0 U; 0 Other;

XX Query Match 63.7%; Score 38.2; DB 12; Length 1378;
XX Best Local Similarity 78.0%; Pred. No. 0.0047;

XX Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
XX
XX 1 GGGTCGACTGCGCCCTCCGGGCGCATGCTGACCTCAGCTCTTGACACAGGCCAGCCCC 59
XX Db 596 GGCTTCACCGCCCGCCCGACGCGGTGTCACCTCGGCCCGGACACACAGGCCGCC 654

RESULT 7

XX ADO23180/c
XX ID ADO23180 standard; RNA; 1424 BP.

XX ADO23180;

XX 12-AUG-2004 (first entry)

XX Antisense human MUC1 mucin glycoprotein RNA (coding sequence) Seqid 75.
XX human; MUC1; mucin glycoprotein; cancer; chemotherapeutic; MUC1/ECB; ss;
XX antisense.

XX Homo sapiens.

XX PN MO200404160-A2.

XX PD 27-MAY-2004.

XX PF 12-NOV-2003; 2003WO-US035848.

XX PR 13-NOV-2002; 2002US-00293391.

XX PR 29-MAY-2003; 2003US-00447839.

XX PA (DAND) DANA FARBER CANCER INST INC.

XX PA (ILEX-) ILEX PROD INC.

XX PI KuFe DW, Kharbada S, Weitman SD;
XX WPI: 2004-420304/39.

XX Double-stranded RNA complex useful for inhibiting proliferation of cancer
PT cell expressing MUC1 mucin glycoprotein, comprises first and second RNA
PT sequences.

XX Disclosure; SEQ ID NO 75; 112pp; English.

XX This invention relates to novel modulators of the human MUC1 mucin
XX glycoprotein for use in cancer therapeutics, where MUC1 is a protein that
XX acts to inhibit the apoptotic response to genotoxic stress caused by
XX chemotherapeutic agents. In particular, it refers to modulators of the
XX MUC1 extracellular domain (MUC1/ECB). The method refers to using double-
XX stranded RNA complexes as MUC1 interference RNA compositions such that
XX MUC1 expression is inhibited, which in turn inhibits cancer cell
XX proliferation. The present invention describes screening assays to
XX identify compounds that inhibit the binding of various MUC1 ligands such
XX as neutregulin 2, as well as agonists, antagonists and antibodies thereof.
XX Furthermore, it provides MUC1 antisense, small interfering RNA (siRNA)
XX and small molecules in combination with chemotherapeutic agents that are
XX useful in the field of cancer therapy. This polynucleotide sequence is
XX the antisense human MUC1 RNA of the invention.

XX Sequence 1424 BP; 309 A; 325 C; 495 G; 0 T; 295 U; 0 Other;

XX Query Match 63.7%; Score 38.2; DB 12; Length 1424;
XX Best Local Similarity 78.0%; Pred. No. 0.0048;

XX Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
XX
XX 1 GGGTCGACTGCGCCCTCCGGGCGCATGCTGACCTCAGCTCTTGACACAGGCCAGCCCC 59
XX Db 1041 GGCTTCACCGCCCGCCCGACGCGGTGTCACCTCGGCCCGGACACACAGGCCGCC 983

RESULT 8

XX ABL60159
XX ID ABL60159 standard; cDNA; 1428 BP.

XX ABL60159;

XX 22-JUL-2002 (first entry)

XX Human MUC1 encoding cDNA SEQ ID NO 2.

XX Human; mucin 1; MUC1; transmembrane protein; SNP; cancer; cytostatic;
XX single nucleotide polymorphism; haplotyping; genotyping; drug;
XX antiinflammatory; gene; ss.

XX Homo sapiens.

XX Key Location/Qualifiers
XX CDS 1..1428

XX /tag=a
XX /product="MUC1"
XX replace(1009,A)
XX /tag=b

XX /standard name="Single nucleotide polymorphism"
XX /note="SNP allelic variation results in Val substituted
XX by Met at position 337 of the MUC1 protein (AB077476)"

XX PN MO200226765-A2.

XX PD 04-APR-2002.

XX PF 25-SEP-2001; 2001WO-US030151.

XX PR 28-SEP-2000; 2000US-0236113P.

XX PA (GENA-) GENAISANCE PHARM INC.

XX PI Chew A, Koshy B;

XX WPI: 2002-405042/43.
XX P-PSDB; ABB77476.

XX New genetic variants of mucin 1, Transmembrane gene, useful in studying
PT expression and function of protein encoded by the gene and for screening
PT drugs to treat diseases e.g. cancer.
XX
PS Claim 23; Fig 2; 75pp; English.
XX
CC The invention relates to a polynucleotide (AB160158, AB160159) encoding
CC mucin 1/MUC1 (AB177476), Transmembrane isogene. The invention describes
CC novel genetic variants of the MUC1 gene. The invention is useful for
CC haplotyping/genotyping the MUC1 gene in an individual and identifying an
CC association between a trait and at least one of the haplotypes or
CC haplotype pairs of MUC1 gene. MUC1 is useful for studying the expression
CC and function of MUC1 and expressing MUC1 protein for use in screening for
CC candidate drugs to treat diseases related to MUC1 activity and in
CC studying the effect of the variation on the biological activity of MUC1
CC as well as on the binding affinity of candidate drugs targeting MUC1 for
CC the treatment of e.g. cancer. MUC1 is further used by the pharmaceutical
CC research scientist to validate MUC1 as a candidate target for and in
CC design of clinical trials of candidate drugs for, treating a specific
CC condition drugs or disease predicted to be associated with MUC1 activity.
CC MUC1 antibodies are useful in a variety of diagnostic and prognostic
CC formats and therapeutic methods
XX
SQ Sequence 1428 BP; 296 A; 496 C; 327 G; 309 T; 0 U; 0 Other;

Query Match 63.7%; Score 38.2; DB 6; Length 1428;
Best Local Similarity 78.0%; Pred. No. 0.0048;
Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 GGGTGACATGCGCCCTCCGGCGCATGCTGACCTCAGCTCCGACACAGAGCCAGCC 59
Db 385 GGCTCACCAGCCGCCGCCACGCCACGCTGTCACCTCCGCCCGACACAGCGGCC 443

RESULT 9
ADO23125
ID ADO23125 standard; RNA; 1428 BP.
XX
AC ADO23125;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human MUC1 mucin glycoprotein RNA (coding sequence) SegID 20.
XX
KW human; MUC1; mucin glycoprotein; cancer; chemotherapeutic; MUC1/ECF; ss.
XX
OS Homo sapiens.
XX
PN WO200404160-A2.
XX
PD 27-MAY-2004.
XX
PF 12-NOV-2003; 2003WO-US035848.
XX
PR 13-NOV-2002; 2002US-00293391.
XX
PR 29-MAY-2003; 2003US-00447839.
XX
PA (DAND) DANA FARBER CANCER INST INC.
XX
PA (ILEX-) ILEX PROD INC.
XX
PI Kuife DW, Kharbada S, Weitman SD;
XX
DR WPI; 2004-420304/39.
XX
PT Double-stranded RNA complex useful for inhibiting proliferation of cancer
PT cell expressing MUC1 mucin glycoprotein, comprises first and second RNA
PT sequences.
XX
PS Claim 2; SEQ ID NO 20; 112pp; English.
XX
CC This invention relates to novel modulators of the human MUC1 mucin
CC glycoprotein for use in cancer therapeutics, where MUC1 is a protein that

CC acts to inhibit the apoptotic response to genotoxic stress caused by
CC chemotherapeutic agents. In particular, the method refers to modulators of the
CC MUC1 extracellular domain (MUC1/ECF). The method refers to using double-
CC stranded RNA complexes as MUC1 interference RNA compositions such that
CC MUC1 expression is inhibited, which in turn inhibits cancer cell
CC proliferation. The present invention describes screening assays to
CC identify compounds that inhibit the binding of various MUC1 ligands such
CC as neuregulin 2, as well as agonists, antagonists and antibodies thereof.
CC Furthermore, it provides MUC1 antisense, small interfering RNA (siRNA)
CC and small molecules in combination with chemotherapeutic agents that are
CC useful in the field of cancer therapy. This polynucleotide sequence is
CC the human MUC1 RNA of the invention.
XX
SQ Sequence 1428 BP; 296 A; 496 C; 327 G; 0 T; 309 U; 0 Other;

Query Match 63.7%; Score 38.2; DB 12; Length 1428;
Best Local Similarity 71.2%; Pred. No. 0.0048;
Matches 42; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

QY 1 GGGTGACATGCGCCCTCCGGCGCATGCTGACCTCAGCTCCGACACAGAGCCAGCC 59
Db 385 GGCTCACCAGCCGCCGCCACGCCACGCTGTCACCTCCGCCCGACACAGCGGCC 443

RESULT 10
ADP32627
ID ADP32627 standard; DNA; 1457 BP.
XX
AC ADP32627;
XX
DT 26-FEB-2004 (first entry)
XX
DE Plasmid JMW358 MUC-1 nucleotide sequence.
XX
KW MUC-1 antigen; immune response; MUC-1; variable number of tandem repeat;
XX VNTR; repeat unit; tumour; metastasis; cytostatic; vaccine; gene therapy;
XX gene; ds.
XX
OS Synthetic.
XX
PN WO2003100060-A2.
XX
PD 04-DEC-2003.
XX
PF 23-MAY-2003; 2003WO-EP005594.
XX
PR 24-MAY-2002; 2002GB-00012046.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI Burden N, Ellis JH, Hamblin PA;
XX
DR WPI; 2004-042811/04.
XX
PT New nucleic acid molecule encoding a MUC-1 antigen, useful for preparing
PT a composition for treating or preventing tumors or metastases.
XX
PS Example; Fig 3; 66pp; English.
XX
CC The present invention describes a nucleic acid molecule which encodes a
CC MUC-1 antigen. The nucleic acid is capable of raising an immune response
CC in vivo, has reduced susceptibility to recombination than full-length MUC
CC -1 and comprises between 1 and 15 variable number of tandem repeats
CC (VNTR) perfect repeat units. Also described: (1) a plasmid comprising the
CC DNA molecule; (2) a protein encoded by the nucleic acid; (3) a
CC pharmaceutical composition comprising the nucleic acid, plasmid or
CC protein and an excipient, diluent or carrier; and (4) a method of
CC treating or preventing tumors or metastases. A MUC1 antigen has
CC cytostatic activity, and can be used in vaccines, and in gene therapy.
CC The nucleic acid is useful for preparing a composition for treating or
CC preventing tumors or metastases. The present sequence is used in the
CC exemplification of the present invention.

SQ Sequence 1457 BP; 302 A; 506 C; 334 G; 315 T; 0 U; 0 Other;

Query Match

Best Local Similarity 63.7%; Score 38.2; DB 12; Length 1457;

Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

DQ 1 GGGTCGACCTGCCCCCGGCGCATGTGTGACCTCCTGACACAGGCCAGCCCC 59
 403 GGGTCGACCTGCCCCCGGCGCATGTGTGACCTCCTGACACAGGCCAGCCCC 461

RESULT 11

AAV48329
 ID AAV48329 standard; cDNA; 1527 BP.

AC AAV48329;

DT 20-NOV-1998 (first entry)

DE MiniMUC1 gene.

KM ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
 KW tumour; tumour-associated antigen.

OS Homo sapiens.

Key Location/Qualifiers
 CDS 1..1527
 /tag= a
 /product= "MiniMUC1 protein"

MO9837095-A2.
 27-AUG-1998.

PF 24-FEB-1998; 98MO-US003693.

PR 24-FEB-1997; 97US-0038253P.

PA (THER-) THERION BIOLOGICS CORP.
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA (DAND) DANA FARRER CANCER INST INC.

PI Schlom J, Kantor J, Kufe D, Pantcali D, Gritz L,

DR WPI; 1998-467492/40.
 DR P-PSDB; AAW77233.

PT New recombinant pox virus for tumour therapy - comprises DNA encoding an
 PT immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
 PS Example 1; Page 21-22; 42pp; English.

CC The immunogenic mini-MUC1 fragment was created from MUC1 tandem repeat
 CC units for inclusion in a recombinant pox virus (RPV). The RPV was used in
 CC a pharmaceutical composition also containing an immunomodulator to
 CC generate MUC1 specific cytotoxic T-lymphocytes. The recombinant pox virus
 CC therefore encodes an immunogenic MUC1 fragment that does not undergo
 CC significant genetic deletion, thereby providing an unexpectedly stable
 CC and immunogenic pox virus. They can be used to prevent or treat tumours
 CC expressing MUC1 tumour-associated antigens

SQ Sequence 1527 BP; 296 A; 573 C; 351 G; 307 T; 0 U; 0 Other;

Query Match

Best Local Similarity 63.7%; Score 38.2; DB 2; Length 1527;

Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

DQ 1 GGGTCGACCTGCCCCCGGCGCATGTGTGACCTCCTGACACAGGCCAGCCCC 59
 226 GGGTCGACCTGCCCCCGGCGCATGTGTGACCTCCTGACACAGGCCAGGCC 284

RESULT 12

AA500585
 ID AA500585 standard; DNA; 1572 BP.

AC AA500585;

DT 12-SEP-2001 (first entry)

DE Human MUC1 DNA sequence.

KM Human; MUC1; antigenic peptide; major histocompatibility complex; MHC-I;
 KW glycoprotein; cytotoxic T lymphocytes; T cell response; cancer; vaccine;
 KW cancer gene therapy; diagnosis; treatment; inflammatory disorder; ds;
 KW organ transplant rejection; graft versus host disease.

OS Homo sapiens.

Key Location/Qualifiers
 CDS 58..1545
 /tag= a
 /product= "Human MUC1"

PN WO200118035-A2.

PD 15-MAR-2001.

PF 07-SEP-2000; 2000WO-EP008761.

PR 08-SEP-1999; 99GB-00021242.

PR 10-SEP-1999; 99EP-00402237.

PR 03-MAR-2000; 2000US-0187215P.

PA (TRGE) TRANSGENE SA.
 PA (IMCR) IMPERIAL CANCER RES TECHNOLOGY LTD.

PI Taylor-Papadimitriou J, Heukamp LC, Offringa R, Melief CJM;

DR WPI; 2001-235187/24.
 DR P-PSDB; AAU00539.

PT New antigenic polypeptides of MUC-1 protein which activate cytotoxic T
 PT lymphocyte proteins and their analogs, useful for identifying a major
 PT histocompatibility complex class I restricted T cell response and for
 PT diagnosing cancer.

PS Claim 1; Page 66-68; 81pp; English.

CC The sequence represents a polynucleotide which encodes MUC1 polypeptide.
 CC derivative antigenic peptides of this protein bind at least one major
 CC histocompatibility complex class I (MHC-I) glycoprotein, which activates
 CC cytotoxic T lymphocytes to induce a protective response against tumours.
 CC diagnosis of cancer involves determining the presence or absence in a
 CC host cell of a MHC class I restricted T cell response to a MUC1
 CC derivative, where the presence of the MHC class I restricted T cell
 CC response indicates that the host has cancer. Measurement of the level of
 CC MHC class I restricted T cell response is also useful to monitor the
 CC severity of cancer, a larger response indicating a more severe cancer.
 CC MUC1 derivatives are useful in cancer therapy and to follow MUC1 specific
 CC immune responses in patients during the course of disease and/or
 CC treatment. MUC1 DNA is useful in cancer gene therapy, vaccination and
 CC diagnostics. Compositions of the sequences are used in vaccines and
 CC treatments against cancer or diseases caused by an immune response, such
 CC as an inflammatory disorder, organ transplant rejection or graft versus
 CC host disease

SQ Sequence 1572 BP; 312 A; 567 C; 364 G; 329 T; 0 U; 0 Other;

Query Match

Best Local Similarity 63.7%; Score 38.2; DB 5; Length 1572;

Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

DQ 1 GGGTCGACCTGCCCCCGGCGCATGTGTGACCTCCTGACACAGGCCAGCCCC 59

Db 442 GGCTCACCAGCCCGCCCGGACCGGTGTACCTCGGCGCCGAGACCAAGCGCGCCCGCC 500

RESULT 13

ADK70370 ID ADK70370 standard; cDNA; 1614 BP.

AC ADK70370;

DT 06-MAY-2004 (first entry)

DE Respiratory disease differentially expressed cDNA #106.

KM ds; gene; cytosolic; respiratory; antiasthmatic; gene therapy;
KM differential gene expression; respiratory disorder; lung cancer;
KM chronic obstructive pulmonary disease; emphysema; asthma.

OS Homo sapiens.

XX WO2003101283-A2.

XX 11-DEC-2003.

XX 02-JUN-2003; 2003WO-US017409.

XX 04-JUN-2002; 2002US-0386005P.

XX (INCY-) INCYTE CORP.

XX Rickett PK, Krasnow R;

XX WPI; 2004-042945/04.

PT New combination comprising cDNAs and proteins that are differentially
PT expressed in respiratory disorders, useful for diagnosing or treating
PT respiratory diseases e.g. lung cancer, chronic obstructive pulmonary
PT diseases or asthma.

PS Claim 1; SEQ ID NO 106; 343bp; English.

CC The invention relates to cDNA sequences that are differentially expressed
CC in respiratory disorders or their complements or encoded proteins. The
CC cDNAs and proteins are useful for diagnosing, treating or monitoring
CC treatment of a subject with a respiratory disease including lung cancer,
CC chronic obstructive pulmonary diseases, emphysema or asthma. The protein
CC is also useful for screening molecules or compounds to identify at least
CC one ligand which specifically binds the protein. It is also useful for
CC preparing and purifying a polyclonal or monoclonal antibody. This
CC sequence corresponds to a cDNA of the invention.

XX Sequence 1614 BP; 329 A; 556 C; 374 G; 353 T; 0 U; 2 Other;

Query Match 63.7%; Score 38.2; DB 12; Length 1614;

Best Local Similarity 78.0%; Pred. No. 0.0048;

Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Oy 1 GGCTGACTGCCCTCCGGCGCATGTGACTGACTGACTGACACCAAGCCGCGCC 59

Db 434 GGCTCACCAGCCCGCCCGGACCGGTGTACCTCGGCGCCGAGACCAAGCGCGCCCGCC 492

RESULT 14

AD157708 ID AD157708 standard; cDNA; 1630 BP.

AC AD157708;

DT 22-APR-2004 (first entry)

DE Human breast specific nucleic acid (BSNA) #79.

KM Human; breast specific nucleic acid; BSNA; gene; ss; metastasis;

KM breast cancer; cytosolic.

XX Homo sapiens.

XX WO2003106648-A2.

XX 24-DEC-2003.

XX 16-JUN-2003; 2003WO-US018934.

XX 14-JUN-2002; 2002US-0389327P.

XX (DIAD-) DIADEXUS INC.

XX Salceda S, Macina RA, Turner LR, Sun Y, Liu C;

XX WPI; 2004-082185/08.

XX P-PSDB; AD157779.

PT Novel isolated polypeptide comprising breast specific protein sequences,
PT useful for diagnosing or monitoring presence and metastases of breast
PT cancer in patient.

PS Claim 1; SEQ ID NO 79; 370pp; English.

CC The invention relates to human breast specific nucleic acids (BSNA) and
CC the breast specific proteins (BSP) they encode. The nucleic acids are
CC useful for determining the presence of a BSNA in a sample which involves
CC contacting the sample with a BSNA under conditions in which the BSNA will
CC selectively hybridise to a BSNA in the sample, and detecting the
CC hybridisation. The nucleic acids are useful for determining the presence
CC of a BSP in a sample which involves contacting the sample with suitable
CC reagent under conditions in which the reagent will selectively interact
CC with the BSP, and detecting the interaction of the reagent with a BSP in
CC the sample. The nucleic acids and proteins are useful for diagnosing or
CC monitoring the presence and metastases of breast cancer in a patient,
CC which involves determining an amount of nucleic acid or protein and
CC comparing the determined amount of nucleic acid or protein in the sample
CC of the patient to the amount of a breast specific marker in a normal
CC control, where a difference in the determined amount in the sample
CC compared to the amount in the control is associated with the presence of
CC breast cancer. The sequences are useful for treating a patient with
CC breast cancer, involving administering a composition consisting of a BSNA
CC or a BSP to a patient, where the administration induces an immune
CC response against the breast cancer cell expressing the BSNA or BSP. This
CC sequence represents a human BSNA of the invention.

XX Sequence 1630 BP; 322 A; 584 C; 405 G; 319 T; 0 U; 0 Other;

Query Match 63.7%; Score 38.2; DB 12; Length 1630;

Best Local Similarity 78.0%; Pred. No. 0.0048;

Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Oy 1 GGCTGACTGCCCTCCGGCGCATGTGACTGACTGACTGACACCAAGCCGCGCC 59

Db 556 GGCTCACCAGCCCGCCCGGACCGGTGTACCTCGGCGCCGAGACCAAGCGCGCCCGCC 654

RESULT 15

AD157689 ID AD157689 standard; cDNA; 1634 BP.

AC AD157689;

DT 22-APR-2004 (first entry)

DE Human breast specific nucleic acid (BSNA) #60.

KM Human; breast specific nucleic acid; BSNA; gene; ss; metastasis;

XX Human; breast specific nucleic acid; BSNA; gene; ss; metastasis;

PN WO2003106648-A2.
 XX
 PD 24-DEC-2003.
 XX
 PF 16-JUN-2003; 2003WO-US018934.
 XX
 PR 14-JUN-2002; 2002US-0389327P.
 XX
 PA (DIAD-) DIADEXUS INC.
 XX
 PI Salceda S, Macina RA, Turner LR, Sun Y, Liu C;
 XX WPI; 2004-082185/08.
 DR P-PSDB; ADI57761.
 XX
 PT Novel isolated polypeptide comprising breast specific protein sequences,
 PT useful for diagnosing or monitoring presence and metastases of breast
 XX cancer in patient.
 PS
 PS Claim 1; SEQ ID NO 60; 370pp; English.
 XX
 CC The invention relates to human breast specific nucleic acids (BSNA) and
 CC the breast specific proteins (BSP) they encode. The nucleic acids are
 CC useful for determining the presence of a BSNA in a sample which involves
 CC contacting the sample with a BSNA under conditions in which the BSNA will
 CC selectively hybridise to a BSNA in the sample, and detecting the
 CC hybridisation. The nucleic acids are useful for determining the presence
 CC of a BSP in a sample which involves contacting the sample with suitable
 CC reagent under conditions in which the reagent will selectively interact
 CC with the BSP, and detecting the interaction of the reagent with a BSP in
 CC the sample. The nucleic acids and proteins are useful for diagnosing or
 CC monitoring the presence and metastases of breast cancer in a patient,
 CC which involves determining amount of nucleic acid or protein in a normal
 CC comparing the determined amount of nucleic acid or protein in the sample
 CC of the patient to the amount of a breast specific marker in a normal
 CC control, where a difference in the determined amount in the sample
 CC compared to the amount in the control is associated with the presence of
 CC breast cancer. The sequences are useful for treating a patient with
 CC breast cancer, involving administering a composition consisting of a BSNA
 CC or a BSP to a patient, where the administration induces an immune
 CC response against the breast cancer cell expressing the BSNA or BSP. This
 CC sequence represents a human BSNA of the invention.
 XX
 SQ Sequence 1634 BP; 349 A; 541 C; 405 G; 337 T; 0 U; 2 Other;
 Query Match 63.7%; Score 38.2; DB 12; Length 1634;
 Best Local Similarity 78.0%; Pred. No. 0.0048;
 Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
 Oy 1 GGGTCGACTGCTCCCTCCGGGCGCATGGTGTGACCTCAGCTCGACACAAAGGCCACCC 59
 Db 596 GGCTCCACCGCCCCCGCCAGCCGCGTGTTCACCTCCGCCCCCGACACAGCGCCGCCCC 654

Search completed: January 15, 2005, 20:36:11
 Job time : 170.5 secs

Query Match 100.0%; Score 60; DB 6; Length 2297;
Best Local Similarity 100.0%; Pred. No. 2.8e-07;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGTCGACCTGCCCCCTCCGGCGCATGTGTGACCTGACCTGACACAGGCCAGCCCCA 60
766 GGGTCGACCTGCCCCCTCCGGCGCATGTGTGACCTGACCTGACACAGGCCAGCCCCA 825

RESULT 2
AX192396 120 bp DNA linear PAT 15-AUG-2001
LOCUS Sequence 2 from Patent EP1103623.
DEFINITION AX192396
ACCESSION AX192396
VERSION AX192396.1 GI:15210363
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 Taylor-Papadimitriou, J., Burchell, J. and Gendler, S.
AUTHORS Human mucin core protein: nucleic acid probes, peptide fragments
TITLE and antibodies thereto, and uses thereof in diagnostic and
JOURNAL therapeutic methods
PATENT: EP 1103623-A 2 30-MAY-2001;
FEATURES IMPERIAL CANCER RESEARCH TECHNOLOGY LIMITED (GB)
Source Location/Qualifiers
1..120
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 63.7%; Score 38.2; DB 6; Length 120;
Best Local Similarity 78.0%; Pred. No. 1.1;
Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 GGGTCGACCTGCCCCCTCCGGCGCATGTGTGACCTGACCTGACACAGGCCAGCCCC 59
87 GGCTCCACCGCCCCCGCCAGCGGTGTCACTCGGCCCGGACACAGGCCGCGCCCC 29

RESULT 3
BD000571/c 120 bp DNA linear PAT 31-JAN-2002
LOCUS Human polymorphic epithelial mucin core protein and nucleic acid
DEFINITION encoding the protein.
ACCESSION BD000571
VERSION BD000571.1 GI:18623684
KEYWORDS JP 2000333675-A/2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 120)
AUTHORS Papadimitriou, J. T., Jendora, S. and Bachieu, J.
TITLE Human polymorphic epithelial mucin core protein and nucleic acid
JOURNAL encoding the protein
PATENT: JP 2000333675-A 2 05-DEC-2000;
COMMENT IMPERIAL CANCER RESEARCH TECHNOLOGY LTD
OS Homo sapiens (human)
PN JP 2000333675-A/2
PD 05-DEC-2000
PR 26-APR-2000 JP 2000125724
PR 07-JAN-1987 GB 8700269, 07-JAN-1987 GB 8700279 PR
22-APR-1987 US 041306, 09-NOV-1987 GB 8726172 FI JOYCE
TAYLOR PAPADIMITRIOU, SANDRA JENORA, JOE BACHIEU PC
C12N5/02, A61K38/00, A61K39/395, A61K49/00, A61P35/00, PC
C07K14/47,
PC C07K16/44, C12N5/10, C12P21/08// (C12N15/02, C12R1:91), (C12N5/10,
C12R1:91),

PC C12N15/00, A61K37/02, C12N5/00, (C12N15/00, C12R1:91), (C12N5/00,
C12R1:91)
CC Key
FH Location/Qualifiers
FT 1..120
FT /organism="Homo sapiens (human)".
FEATURES
source Location/Qualifiers
1..120
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 63.7%; Score 38.2; DB 6; Length 120;
Best Local Similarity 78.0%; Pred. No. 1.1;
Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 GGGTCGACCTGCCCCCTCCGGCGCATGTGTGACCTGACCTGACACAGGCCAGCCCC 59
87 GGCTCCACCGCCCCCGCCAGCGGTGTCACTCGGCCCGGACACAGGCCGCGCCCC 29

RESULT 4
AF423031 1414 bp mRNA linear SYN 10-JUL-2003
LOCUS Synthetic construct Homo sapiens mucin variant MUC1-CT58 (MUC1)
DEFINITION AF423031
ACCESSION AF423031
VERSION AF423031.1 GI:19338621
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1 (bases 1 to 1414)
AUTHORS Hinojosa-Kurtzberg, A.M., Johansson, M.E., Madsen, C.S., Hansson, G.C.
TITLE Novel MUC1 splice variants contribute to mucin overexpression in
JOURNAL CPTFR-deficient mice
Am. J. Physiol. Gastrointest. Liver Physiol. 284 (5), G853-G862
(2003)
MEDLINE 22570517
PUBMED 12529261
2 (bases 1 to 1414)
Hinojosa-Kurtzberg, A.M. and Gendler, S.J.
AUTHORS Direct Submission
TITLE Submitted (21-SEP-2001) Biochemistry, Mayo Clinic, 13400 E. Shea
JOURNAL Boulevard, Scottsdale, AZ 85259, USA
FEATURES
source Location/Qualifiers
1..1414
/organism="synthetic construct"
/mol_type="mRNA"
/db_xref="taxon:32630"
/note="Homo sapiens gene in transgenic Mus musculus
C57BL/6; isolated from intestinal mucosa"
1..1414
/gene="MUC1"
1..1386
/gene="MUC1"
/note="alternatively spliced; contains exon 6b resulting
in variant carboxy-terminal domain; lacks sites for
beta-catenin and Grb2 interactions; derived from Homo
sapiens"
/codon_start=1
/transl_table=11
/product="mucin variant MUC1-CT58"
/protein_id="AA186735.1"
/db_xref="GI:19338622"
/translation="MTPGQSPFPLLLTLVTVVYSGHSGSPGEGKETSATGORS
VPSSTKNAVSMSTSVLSHSGSGSTTGQDVLAPATPPASGSAATWQDVTYSV
VIRPLAGSTTPPAHVTSAPDKPAAGSTAPPAHVTSAPTPPARPGSTAPPAAGVY
ABDNPAAGSTAPPAHVTSAGSASGASSTVHNGTSARATTPPASSTFSPSH
SPTFTLASHSTKTDASTHSTVPLTSSNHSSTSPOLSTGVGFFLSFHLSTNQFNS
SLSDPSTDYKQELRDISMFLQIKQGGFLGLSNIKRPGSVVVQTLTARREGTINV

HDVETQFNQKTEAASRYNLITSDVSPDPFSAQSGAVPGWGLALVLVCLVA
LAIVYLIALAVCGGRKNVYGGDIFPARDTYHPMBEYPTVYHNGYVPPSSIDRSBYE
KEMVDRERLRA"

ORIGIN

Query Match 63.7%; Score 38.2; DB 12; Length 1414;
Best Local Similarity 78.0%; Pred. No. 0.59;
Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 1 GGGTGACTGCCCCCTCCGGCGCATGTGTGACCTCAGCTCTGACACAAAGCCAGCCCC 59
Db 385 GGCTCACCAGCCCCCGCCAGCCACGCGTGCACCTCGGCCCGGACACAGCGCGGCC 443

RESULT 5
CQ715242 1455 bp DNA linear PAT 03-FEB-2004
LOCUS Sequence 1176 from Patent WO02068579.
ACCESSION CQ715242
VERSION CQ715242.1 GI:42276099
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1
AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
TITLE Kites, such as nucleic acid arrays, comprising a majority of humenexons or transcripts, for detecting expression and other uses thereof
JOURNAL Patent: WO 02068579-A 1176 06-SEP-2002;
PE Corporation (NY) (US)

FEATURES
source 1.1455
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 63.7%; Score 38.2; DB 6; Length 1455;
Best Local Similarity 78.0%; Pred. No. 0.59;
Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 1 GGGTGACTGCCCCCTCCGGCGCATGTGTGACCTCAGCTCTGACACAAAGCCAGCCCC 59
Db 412 GGCTCACCAGCCCCCGCCAGCCACGCGTGCACCTCGGCCCGGACACAGCGCGGCC 470

RESULT 6
AX959914 1457 bp DNA linear PAT 14-JAN-2004
LOCUS Sequence 19 from Patent WO0310060.
ACCESSION AX959914
VERSION AX959914.1 GI:40880143
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1
AUTHORS Burden, N.G., Ellis, J.H. and Hamblin, P.A.
TITLE Muc-1 antigen with reduced number of vntc repeat units
JOURNAL Patent: WO 0310060-A 19 04-DEC-2003;
GLAXO GROUP LIMITED (GB)

FEATURES
source 1.1457
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 63.7%; Score 38.2; DB 6; Length 1457;
Best Local Similarity 78.0%; Pred. No. 0.59;

Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
Qy 1 GGGTGACTGCCCCCTCCGGCGCATGTGTGACCTCAGCTCTGACACAAAGCCAGCCCC 59
Db 403 GGCTCACCAGCCCCCGCCAGCCACGCGTGCACCTCGGCCCGGACACAGCGCGGCC 461

RESULT 7
AX093798 1572 bp DNA linear PAT 30-MAR-2001
LOCUS Sequence 1 from Patent WO0118035.
ACCESSION AX093798
VERSION AX093798.1 GI:13510057
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1
AUTHORS Taylor, Papadimitriou, J., Heukamp, L.C., Offringa, R., Melief, C.J., Acres, B. and Thomas, M.
TITLE Muc-1 derived peptides
JOURNAL Patent: WO 0118035-A 1 15-MAR-2001;
TRANSCENE S.A. (FR) ; IMPERIAL CANCER RESEARCH TECHNOLOGY LIMITED (GB)

FEATURES
source 1.1572
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
CDS 58..1545
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC35422.1"
/translation="MTPTGSPFFLLTLVTVTGSGHASSTPGCEKETSATORSS
VPSSTKNAVSWTSVLSHSGSSTTGODVLAATBPASGSAATWGDVTSVP
VTRPALGSTTPPAHVTSAAPDNKPAKPGSTAPPAHVTSAAPDPGSAAPAHVTS
APDTPPAGSTAPPAHVTSAAPDNKPAKPGSTAPPAHVTSAAPDPGSAAPAHVTS
ARATTPPAKSTSPSPISPHSHSDPTTLASHSTKIDASTHSTVPLTSSNHSSTPQL
STGVSFPLSPHISNLQPNSSLSBPDSTDYDELQDISEMPLOIKOGGFLGSLNKF
RGSVVOLTLAFTREGTINVDVETQFNQKTEAASRYNLITSDVSPDPFSAQSGAVPGWGLALVLCVLA
VCGGRKNVYGGDIFPARDTYHPMBEYPTVYHNGYVPPSSIDRSBYE
TYHTRGRVPPSSIDRSBYEKSAGNGSSLSYTPNPAVAATSAANL"

ORIGIN

Query Match 63.7%; Score 38.2; DB 6; Length 1572;
Best Local Similarity 78.0%; Pred. No. 0.57;
Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 1 GGGTGACTGCCCCCTCCGGCGCATGTGTGACCTCAGCTCTGACACAAAGCCAGCCCC 59
Db 442 GGCTCACCAGCCCCCGCCAGCCACGCGTGCACCTCGGCCCGGACACAGCGCGGCC 500

RESULT 8
AX959684 1774 bp DNA linear PAT 14-JAN-2004
LOCUS Sequence 10 from Patent WO03099193.
ACCESSION AX959684
VERSION AX959684.1 GI:40880030
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1
AUTHORS Burden, N. and Hamblin, P.
TITLE Vaccines
JOURNAL Patent: WO 03099193-A 10 04-DEC-2003;
GLAXO GROUP LIMITED (GB)

FEATURES
source 1.1774
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

Query Match 63.7%; Score 38.2; DB 6; Length 1774;
Best Local Similarity 78.0%; Pred. No. 0.56;
Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 GGGTCGACTGCCCCCTCCGGCGCATGTGACCTCAGCTCCTGACACAGGCCAGCCCC 59
462 GGCTCCACCGCCCCCCAGCCACGAGTGTACCTCGCCCCGGACACAGGCCAGCCCC 520

RESULT 9
AX959912 1774 bp DNA linear PAT 14-JAN-2004
LOCUS Sequence 17 from Patent WO03100060.
DEFINITION AX959912
ACCESSION AX959912 GI:40880142
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1
AUTHORS Burden,N.G., Ellis,J.H. and Hamblin,P.A.
TITLE Muc-1 antigen with reduced number of vntir repeat units
JOURNAL Patent: WO 03100060-A, 17 04-DEC-2003;
GLAXO GROUP LIMITED (GB)

FEATURES
source Location/Qualifiers
1..1774

ORIGIN
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

Query Match 63.7%; Score 38.2; DB 6; Length 1774;
Best Local Similarity 78.0%; Pred. No. 0.56;
Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 GGGTCGACTGCCCCCTCCGGCGCATGTGACCTCAGCTCCTGACACAGGCCAGCCCC 59
Db 462 GGCTCCACCGCCCCCCAGCCACGAGTGTACCTCGCCCCGGACACAGGCCAGCCCC 520

RESULT 10
AR492318 1804 bp DNA linear PAT 15-MAY-2004
LOCUS Sequence 17 from patent US 6716627.
DEFINITION AR492318
ACCESSION AR492318
VERSION AR492318.1 GI:47260892
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1804)
AUTHORS Dobie,K.W.
TITLE Antisense modulation of mucin 1, transmembrane expression
JOURNAL Patent: US 6716627-A 17 06-APR-2004;
FEATURES Location/Qualifiers
1..1804
source /organism="unknown"
/mol_type="genomic DNA"

ORIGIN
/organism="unknown"
/mol_type="genomic DNA"

Query Match 63.7%; Score 38.2; DB 6; Length 1804;
Best Local Similarity 78.0%; Pred. No. 0.56;
Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 GGGTCGACTGCCCCCTCCGGCGCATGTGACCTCAGCTCCTGACACAGGCCAGCCCC 59
Db 457 GGCTCCACCGCCCCCCAGCCACGAGTGTACCTCGCCCCGGACACAGGCCAGCCCC 515

RESULT 11
AX335367 1804 bp DNA linear PAT 09-JAN-2002
LOCUS Sequence 5876 from Patent WO0194629.
DEFINITION AX335367
ACCESSION AX335367
VERSION AX335367.1 GI:18126086
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1
AUTHORS Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
Hortigan,S., Soppet,D.R. and Weaver,Z.
TITLE Cancer gene determination and therapeutic screening using signature
JOURNAL gene sets
Patent: WO 0194629-A 5876 13-DEC-2001;
Avalon Pharmaceuticals (US)

FEATURES
source Location/Qualifiers
1..1804
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 63.7%; Score 38.2; DB 6; Length 1804;
Best Local Similarity 78.0%; Pred. No. 0.56;
Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 GGGTCGACTGCCCCCTCCGGCGCATGTGACCTCAGCTCCTGACACAGGCCAGCCCC 59
Db 457 GGCTCCACCGCCCCCCAGCCACGAGTGTACCTCGCCCCGGACACAGGCCAGCCCC 515

RESULT 12
HMMUCAB 1804 bp mRNA linear PRI 07-JAN-1995
LOCUS Human polymorphic epithelial mucin (PEM) mRNA, complete cds.
DEFINITION J05581
ACCESSION J05581 GI:188869
VERSION J05581.1 GI:188869
KEYWORDS polymorphic epithelial mucin.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 1804)
AUTHORS Gendler,S.J., Lancaster,C.A., Taylor-Papadimitriou,J., Duhig,T.,
Peat,N., Burchell,J., Pemberton,L., Ialenti,E.N. and Wilson,D.
TITLE Molecular cloning and expression of human tumor-associated
JOURNAL polymorphic epithelial mucin
J. Biol. Chem. 265 (25), 15286-15293 (1990)
MEDLINE 90368715
PUBMED 1697589

COMMENT Original source text: Homo sapiens adult adenocarcinoma CDNA to
Draft entry and computer-readable sequence for [J. Biol. Chem.
(1990) in press] kindly submitted
by S.J. Gendler, 26-JUN-1990.
FEATURES Location/Qualifiers
1..1804
source /organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/map="Xq26.3-q27.1"
/cell_line="BT20"
/tissue_type="adenocarcinoma"
/dev stage="adult"
1..1804
/gene="MCNNA"
/gene="MCNNA"
/gene="MCNNA"

gene
mRNA

GLAXO GROUP LIMITED (GB)
 FEATURES Location/Qualifiers
 source 1..2135
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN

Query Match 63.7%; Score 38.2; DB 6; Length 2135;
 Best Local Similarity 78.0%; Pred.No. 0.53;
 Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 GGATGACTGCCCCCTCCGGCGCATGATGACCTCAGCTCCTTGAACAAGGCCAGCCCC 59
 |||||
 Db 462 GGCTCCACCGCCCCCCCGAGCCACGCGTGTCACTCGGCCCGAGACCCAGCGCGGCCCC 520
 |||||

Search completed: January 15, 2005, 22:36:54
 Job time : 723.8 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2005, 22:37:00 : Search time 183.4 Seconds
(without alignments)
1879.789 Million cell updates/sec

Title: US-10-057-136-9

Perfect score: 60

Sequence: 1 ggaagtagccgcacccacccgc.....cagacaccgcactgcgcca 60

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 4300275 seqs, 2872944193 residues

Total number of hits satisfying chosen parameters: 8600550

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications NA.*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	60	100.0	60	US-10-057-136-9	Sequence 9, Appli
2	60	100.0	2297	US-10-406-317-41	Sequence 41, Appli
3	52	86.7	60	US-10-057-136-8	Sequence 8, Appli
4	38.2	63.7	60	US-10-057-136-12	Sequence 12, Appli
5	33.4	55.7	60	US-10-057-136-4	Sequence 4, Appli
6	31.8	53.0	60	US-10-057-136-5	Sequence 5, Appli
7	31.4	52.3	93	US-10-471-607-3	Sequence 3, Appli
8	31.4	52.3	156	US-10-471-607-5	Sequence 5, Appli
9	31.4	52.3	157	US-10-471-607-6	Sequence 6, Appli
10	31.4	52.3	157	US-10-471-607-9	Sequence 9, Appli
11	31.4	52.3	3621	US-10-471-607-11	Sequence 11, Appli
12	31.4	52.3	6877	US-10-471-607-13	Sequence 13, Appli

13	30	50.0	60	US-10-057-136-6	Sequence 6, Appli
14	29	48.3	60	US-10-057-136-7	Sequence 7, Appli
15	29	48.3	60	US-10-057-136-11	Sequence 11, Appli
16	28	46.7	1845	US-10-156-761-3619	Sequence 3619, Ap
17	28	45.7	9025608	US-10-156-761-1	Sequence 1, Appli
18	27.4	45.7	60	US-10-057-136-2	Sequence 2, Appli
19	27.4	45.7	1424	US-10-447-839A-75	Sequence 75, Appli
20	27.4	45.7	1428	US-10-447-839A-20	Sequence 20, Appli
21	27.4	45.7	1527	US-10-057-136-19	Sequence 19, Appli
22	27.4	45.7	1799	US-10-447-839A-19	Sequence 19, Appli
23	27.4	45.7	1804	US-09-964-824A-573	Sequence 573, App
24	27.4	45.7	1804	US-10-029-517-17	Sequence 17, Appli
25	27.4	45.7	1804	US-10-721-597-30	Sequence 30, Appli
26	27.4	45.7	1804	US-10-775-920-84	Sequence 84, Appli
27	27.4	45.7	1823	US-10-101-510-339	Sequence 339, App
28	27.4	45.7	4139	US-09-964-824A-105	Sequence 105, App
29	27.4	45.7	4139	US-09-964-824A-578	Sequence 578, App
30	27.4	45.7	4139	US-09-864-864-334	Sequence 334, App
31	27.4	45.7	4139	US-09-880-107-2121	Sequence 2121, Ap
32	27.4	45.7	4139	US-09-968-007A-751	Sequence 751, App
33	27.4	45.7	4139	US-10-171-311-157	Sequence 157, App
34	27.4	45.7	4139	US-10-177-231-310	Sequence 310, App
35	27.4	45.7	4139	US-10-440-464-155	Sequence 155, App
36	27.4	45.7	4139	US-10-734-564-53	Sequence 53, Appli
37	27.4	45.7	4139	US-10-775-920-80	Sequence 80, Appli
38	27.4	45.7	4139	US-10-775-920-85	Sequence 85, Appli
39	27.4	45.7	8181	US-10-447-839A-18	Sequence 18, Appli
40	27.4	45.7	8186	US-10-247-703-23	Sequence 23, Appli
41	27.4	45.7	8186	US-10-029-517-19	Sequence 19, Appli
42	26.2	43.7	2649	US-10-369-493-42923	Sequence 42923, A
43	26	43.3	518	US-10-247-703-38	Sequence 38, Appli
44	26	43.3	518	US-10-029-517-101	Sequence 101, App
45	25.8	43.0	78	US-10-057-136-13	Sequence 13, Appli

ALIGNMENTS

RESULT 1
US-10-057-136-9
; Sequence 9, Application US/10057136
; Publication No. US20030021770A1
; GENERAL INFORMATION:
; APPLICANT: SCHLON, JEFFREY
; APPLICANT: KANTOR, JUDITH
; APPLICANT: KUFE, DONALD
; APPLICANT: PANICALI, DENNIS
; APPLICANT: GRITZ, LINDA
; TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
; FILE REFERENCE: 700953/47113C
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/366, 670
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: PCT/US98/03693
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/038, 253
; PRIOR FILING DATE: 1997-02-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-136-9

Query Match 100.0%; Score 60; DB 14; Length 60;
Best Local Similarity 100.0%; Pred. No. 1.9e-12;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 GGAAGTAGCCGCACCCACCCGCAGGAGTCAACAGCGCCGACACTGCAGCTGCAGCA 60
|||||

Db

1 GGAAGTACCGCTCCACTGTCACACGGGGGTCAAGAGCGCGCCAGACACTGACTGGGCCA 60

RESULT 2

US-10-406-317-41

; Sequence 41, Application US/10406317
; Publication No. US20040019195A1
; GENERAL INFORMATION:

; APPLICANT: Schlom, Jeffrey;

; APPLICANT: Hodge, James;

; APPLICANT: Panicali, Dennis

; TITLE OF INVENTION: A recombinant vector expressing multiple constimulatory

; FILE REFERENCE: 38163-0189

; CURRENT APPLICATION NUMBER: US/10/406,317

; PRIOR FILING DATE: 2003-04-04

; PRIOR FILING DATE: 2001-05-30

; PRIOR FILING DATE: 1999-11-12

; PRIOR FILING DATE: 1998-12-09

; NUMBER OF SEQ ID NOS: 41

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 41

; LENGTH: 2297

; TYPE: DNA

; ORGANISM: Artificial Sequence

; OTHER INFORMATION: Description of Artificial Sequence: VECTOR

US-10-406-317-41

Query Match

Best Local Similarity 100.0%; Score 60; DB 16; Length 2297;

Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGAAGTACCGCTCCACTGTCACACGGGGGTCAAGAGCGCGCCAGACACTGACTGGGCCA 60

Db 706 GGAAGTACCGCTCCACTGTCACACGGGGGTCAAGAGCGCGCCAGACACTGACTGGGCCA 765

RESULT 3

US-10-057-136-8

; Sequence 8, Application US/10057136
; Publication No. US20030021770A1
; GENERAL INFORMATION:

; APPLICANT: SCHLOM, JEFFREY

; APPLICANT: KANTOR, JUDITH

; APPLICANT: KUFE, DONALD

; APPLICANT: PANICALI, DENNIS

; APPLICANT: GRITZ, LINDA

; TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1

; FILE REFERENCE: 700953/47113C

; CURRENT APPLICATION NUMBER: US/10/057,136

; PRIOR FILING DATE: 2002-01-25

; PRIOR FILING DATE: 1999-09-03

; PRIOR FILING DATE: 1998-02-24

; PRIOR FILING DATE: 1997-02-24

; NUMBER OF SEQ ID NOS: 20

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 8

; LENGTH: 60

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-057-136-8

Query Match

Best Local Similarity 86.7%; Score 52; DB 14; Length 60;

US-10-057-136-8

Matches 55; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 GGAAGTACCGCTCCACTGTCACACGGGGGTCAAGAGCGCGCCAGACACTGACTGGGCCA 60

Db 1 GGAAGTACCGCTCCACTGTCACACGGGGGTCAAGAGCGCGCCAGACACTGACTGGGCCA 60

RESULT 4

US-10-057-136-12

; Sequence 12, Application US/10057136
; Publication No. US20030021770A1
; GENERAL INFORMATION:

; APPLICANT: SCHLOM, JEFFREY

; APPLICANT: KANTOR, JUDITH

; APPLICANT: KUFE, DONALD

; APPLICANT: PANICALI, DENNIS

; APPLICANT: GRITZ, LINDA

; TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1

; FILE REFERENCE: 700953/47113C

; CURRENT APPLICATION NUMBER: US/10/057,136

; PRIOR FILING DATE: 2002-01-25

; PRIOR FILING DATE: 1999-08-03

; PRIOR FILING DATE: 1998-02-24

; PRIOR FILING DATE: 1997-02-24

; NUMBER OF SEQ ID NOS: 20

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 12

; LENGTH: 60

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-057-136-12

Query Match

Best Local Similarity 63.7%; Score 38.2; DB 14; Length 60;

Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 1 GGAAGTACCGCTCCACTGTCACACGGGGGTCAAGAGCGCGCCAGACACTGACTGGGCC 59

Db 1 GGAAGTACCGCTCCACTGTCACACGGGGGTCAAGAGCGCGCCAGACACTGACTGGGCC 59

RESULT 5

US-10-057-136-4

; Sequence 4, Application US/10057136
; Publication No. US20030021770A1
; GENERAL INFORMATION:

; APPLICANT: SCHLOM, JEFFREY

; APPLICANT: KANTOR, JUDITH

; APPLICANT: KUFE, DONALD

; APPLICANT: PANICALI, DENNIS

; APPLICANT: GRITZ, LINDA

; TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1

; FILE REFERENCE: 700953/47113C

; CURRENT APPLICATION NUMBER: US/10/057,136

; PRIOR FILING DATE: 2002-01-25

; PRIOR FILING DATE: 1999-08-03

; PRIOR FILING DATE: 1998-02-24

; PRIOR FILING DATE: 1997-02-24

; NUMBER OF SEQ ID NOS: 20

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 4

; LENGTH: 60

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-057-136-4

Query Match 55.7%; Score 33.4; DB 14; Length 60;
Best Local Similarity 72.9%; Pred. No. 0.011;
Matches 43; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 1 GGAAGTACCGCTCCACCTGACACAGCGGGTCAAGAGCCGACACTGAGCTGGGCC 59
Db 1 GGCACTACTGCACACCGGCACATGGCGTAACTACGACACTGATACAGACTCTGCACC 59

RESULT 6
US-10-057-136-5

Sequence 5, Application US/10057136
Publication No. US20030021770A1

GENERAL INFORMATION:

APPLICANT: SCHLOM, JEFFREY

APPLICANT: KANTOR, JUDITH

APPLICANT: KUFE, DONALD

APPLICANT: PANICALI, DENNIS

APPLICANT: GRITZ, LINDA

TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1

FILE REFERENCE: 700953/47113C

CURRENT APPLICATION NUMBER: US/10/057,136

PRIOR FILING DATE: 2002-01-25

PRIOR APPLICATION NUMBER: 09/366,670

PRIOR FILING DATE: 1999-08-03

PRIOR APPLICATION NUMBER: PCT/US98/03693

PRIOR FILING DATE: 1998-02-24

PRIOR APPLICATION NUMBER: 60/038,253

PRIOR FILING DATE: 1997-02-24

NUMBER OF SEQ ID NOS: 20

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 5

LENGTH: 60

TYPE: DNA

ORGANISM: Homo sapiens

US-10-057-136-5

Query Match 53.0%; Score 31.8; DB 14; Length 60;
Best Local Similarity 71.2%; Pred. No. 0.042;
Matches 42; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 1 GGAAGTACCGCTCCACCTGACACAGCGGGTCAAGAGCCGACACTGAGCTGGGCC 59
Db 1 GGATCACCGCGCTGCGCGACGAGTGAAGTCAAGTCAAGAGCTGGGCCGCTCC 59

RESULT 7
US-10-471-607-3

Sequence 3, Application US/10471607
Publication No. US20040115740A1

GENERAL INFORMATION:

APPLICANT: The Victoria University of Manchester

APPLICANT: Benson, Roderick

TITLE OF INVENTION: Intracellular analysis.

FILE REFERENCE: P088857PWO

CURRENT APPLICATION NUMBER: US/10/471,607

CURRENT FILING DATE: 2003-09-24

PRIOR APPLICATION NUMBER: GB 0108165.2

PRIOR FILING DATE: 3001-03-21

NUMBER OF SEQ ID NOS: 16

SOFTWARE: PatentIn version 3.1

SEQ ID NO 3

LENGTH: 93

TYPE: DNA

ORGANISM: Artificial

FEATURE:

OTHER INFORMATION: Artificial epitope construct

US-10-471-607-3

Query Match 52.3%; Score 31.4; DB 17; Length 93;
Best Local Similarity 85.4%; Pred. No. 0.058;

Matches 35; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 GGAAGTACCGCTCCACCTGACACAGCGGGTCAAGAGCCGACACTGAGCTGGGCC 41
Db 38 GGATCTACCGCTCTCTGCGCGACGAGTCAAGAGCTGGGCCGCTCC 78

RESULT 8
US-10-471-607-5

Sequence 5, Application US/10471607
Publication No. US20040115740A1

GENERAL INFORMATION:

APPLICANT: The Victoria University of Manchester

APPLICANT: Benson, Roderick

TITLE OF INVENTION: Intracellular analysis.

FILE REFERENCE: P088857PWO

CURRENT APPLICATION NUMBER: US/10/471,607

CURRENT FILING DATE: 2003-09-24

PRIOR APPLICATION NUMBER: GB 0108165.2

PRIOR FILING DATE: 3001-03-21

NUMBER OF SEQ ID NOS: 16

SOFTWARE: PatentIn version 3.1

SEQ ID NO 5

LENGTH: 156

TYPE: DNA

ORGANISM: Artificial

FEATURE:

OTHER INFORMATION: Artificial epitope construct

US-10-471-607-5

Query Match 52.3%; Score 31.4; DB 17; Length 156;
Best Local Similarity 85.4%; Pred. No. 0.058;
Matches 35; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 GGAAGTACCGCTCCACCTGACACAGCGGGTCAAGAGCCGACACTGAGCTGGGCC 41
Db 38 GGATCTACCGCTCTCTGCGCGACGAGTCAAGAGCTGGGCCGCTCC 78

RESULT 9
US-10-471-607-6

Sequence 6, Application US/10471607
Publication No. US20040115740A1

GENERAL INFORMATION:

APPLICANT: The Victoria University of Manchester

APPLICANT: Benson, Roderick

TITLE OF INVENTION: Intracellular analysis.

FILE REFERENCE: P088857PWO

CURRENT APPLICATION NUMBER: US/10/471,607

CURRENT FILING DATE: 2003-09-24

PRIOR APPLICATION NUMBER: GB 0108165.2

PRIOR FILING DATE: 3001-03-21

NUMBER OF SEQ ID NOS: 16

SOFTWARE: PatentIn version 3.1

SEQ ID NO 6

LENGTH: 157

TYPE: DNA

ORGANISM: Artificial

FEATURE:

OTHER INFORMATION: Artificial epitope construct

US-10-471-607-6

Query Match 52.3%; Score 31.4; DB 17; Length 157;
Best Local Similarity 85.4%; Pred. No. 0.058;
Matches 35; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 GGAAGTACCGCTCCACCTGACACAGCGGGTCAAGAGCCGACACTGAGCTGGGCC 41
Db 38 GGATCTACCGCTCTCTGCGCGACGAGTCAAGAGCTGGGCCGCTCC 78

RESULT 10
US-10-471-607-9

```
/ Sequence 9, Application US/10471607
/ Publication No. US20040115740A1
/ GENERAL INFORMATION:
/ APPLICANT: The Victoria University of Manchester
/ APPLICANT: Benson, Roderick
/ TITLE OF INVENTION: Intracellular analysis.
/ FILE REFERENCE: P088857PMO
/ CURRENT APPLICATION NUMBER: US/10/471,607
/ PRIOR FILING DATE: 2003-09-24
/ PRIOR APPLICATION NUMBER: GB 0108165.2
/ NUMBER OF SEQ ID NOS: 16
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 9
/ LENGTH: 157
/ TYPE: DNA
/ ORGANISM: Artificial
/ FEATURE:
/ OTHER INFORMATION: Artificial epitope construct
US-10-471-607-9
```

Query Match 52.3%; Score 31.4; DB 17; Length 157;
Best Local Similarity 85.4%; Pred. No. 0.058; Mismatches 6; Indels 0; Gaps 0;

Matches 35; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GGAAGTACCGCTCTCCACGCGGGTCAACAAGCGCGCC 41
DB 38 GGAATCTACCGCTCTCTCCACGCGGGTCAACAAGCGCAC 78

```
RESULT 11
US-10-471-607-11
/ Sequence 11, Application US/10471607
/ Publication No. US20040115740A1
/ GENERAL INFORMATION:
/ APPLICANT: The Victoria University of Manchester
/ APPLICANT: Benson, Roderick
/ TITLE OF INVENTION: Intracellular analysis.
/ FILE REFERENCE: P088857PMO
/ CURRENT APPLICATION NUMBER: US/10/471,607
/ PRIOR FILING DATE: 2003-09-24
/ PRIOR APPLICATION NUMBER: GB 0108165.2
/ NUMBER OF SEQ ID NOS: 16
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 11
/ LENGTH: 3621
/ TYPE: DNA
/ ORGANISM: Artificial
/ FEATURE:
/ OTHER INFORMATION: Engineered construct
US-10-471-607-11
```

Query Match 52.3%; Score 31.4; DB 17; Length 3621;
Best Local Similarity 85.4%; Pred. No. 0.055; Mismatches 6; Indels 0; Gaps 0;

Matches 35; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GGAAGTACCGCTCTCCACGCGGGTCAACAAGCGCGCC 41
DB 659 GGAATCTACCGCTCTCTCCACGCGGGTCAACAAGCGCAC 699

```
RESULT 12
US-10-471-607-13
/ Sequence 13, Application US/10471607
/ Publication No. US20040115740A1
/ GENERAL INFORMATION:
/ APPLICANT: The Victoria University of Manchester
/ APPLICANT: Benson, Roderick
/ TITLE OF INVENTION: Intracellular analysis.
/ FILE REFERENCE: P088857PMO
/ CURRENT APPLICATION NUMBER: US/10/471,607
/ CURRENT FILING DATE: 2003-09-24
```

```
/ PRIOR APPLICATION NUMBER: GB 0108165.2
/ PRIOR FILING DATE: 3001-03-21
/ NUMBER OF SEQ ID NOS: 16
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 13
/ LENGTH: 6877
/ TYPE: DNA
/ ORGANISM: Artificial
/ FEATURE:
/ OTHER INFORMATION: Engineered construct
US-10-471-607-13
```

Query Match 52.3%; Score 31.4; DB 17; Length 6877;
Best Local Similarity 85.4%; Pred. No. 0.054; Mismatches 6; Indels 0; Gaps 0;

Matches 35; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GGAAGTACCGCTCTCCACGCGGGTCAACAAGCGCGCC 41
DB 701 GGAATCTACCGCTCTCTCCACGCGGGTCAACAAGCGCAC 741

```
RESULT 13
US-10-057-136-6
/ Sequence 6, Application US/10057136
/ Publication No. US20030021770A1
/ GENERAL INFORMATION:
/ APPLICANT: SCHLOM, JEFFREY
/ APPLICANT: KANTOR, JUDITH
/ APPLICANT: KUPF, DONALD
/ APPLICANT: PANICALI, DENNIS
/ TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
/ FILE REFERENCE: 700953/47113C
/ CURRENT APPLICATION NUMBER: US/10/057,136
/ PRIOR FILING DATE: 2002-01-25
/ PRIOR APPLICATION NUMBER: 09/366,670
/ PRIOR FILING DATE: 1999-08-03
/ PRIOR APPLICATION NUMBER: PCT/US98/03693
/ PRIOR FILING DATE: 1998-02-24
/ PRIOR APPLICATION NUMBER: 60/038,253
/ NUMBER OF SEQ ID NOS: 20
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 6
/ LENGTH: 60
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-057-136-6
```

Query Match 50.0%; Score 30; DB 14; Length 60;
Best Local Similarity 72.2%; Pred. No. 0.19; Mismatches 15; Indels 0; Gaps 0;

Matches 39; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 7 ACCGCTCACCTGACACGCGGGTCAACAAGCGCGCGACACTGCTGCGCA 60
DB 7 ACACTCTCTCCCGCTCATGGGTTACTTCTGCTCCAGATCTGCGCGCACTCA 60

```
RESULT 14
US-10-057-136-7
/ Sequence 7, Application US/10057136
/ Publication No. US20030021770A1
/ GENERAL INFORMATION:
/ APPLICANT: SCHLOM, JEFFREY
/ APPLICANT: KANTOR, JUDITH
/ APPLICANT: KUPF, DONALD
/ APPLICANT: PANICALI, DENNIS
/ TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
/ FILE REFERENCE: 700953/47113C
/ CURRENT APPLICATION NUMBER: US/10/057,136
```

/ CURRENT FILING DATE: 2002-01-25
/ PRIOR APPLICATION NUMBER: 09/366,670
/ PRIOR FILING DATE: 1999-08-03
/ PRIOR APPLICATION NUMBER: PCT/US98/03693
/ PRIOR FILING DATE: 1998-02-24
/ PRIOR APPLICATION NUMBER: 60/038,253
/ PRIOR FILING DATE: 1997-02-24
/ NUMBER OF SEQ ID NOS: 20
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO: 7
/ LENGTH: 60
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-057-136-7

Query Match 48.3%; Score 29; DB 14; Length 60;
Best Local Similarity 71.7%; Pred. No. 0.45;
Matches 38; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 7 ACCGCTCCACCTGCACAGCGGGGTCAACAGCGCCGACACTGACCTGGCC 59
DB 7 ACGGCCCCCTGCTCAGCGGTGTACATCCGCCCGGATATCCAGACCGGCCCC 59

RESULT 15
US-10-057-136-11

/ Sequence 11, Application US/10057136
/ Publication No. US2003002170A1
/ GENERAL INFORMATION:
/ APPLICANT: SCHLOM, JEFFREY
/ APPLICANT: KANTOR, JUDITH
/ APPLICANT: KURE, DONALD
/ APPLICANT: PANICALI, DENNIS
/ APPLICANT: GRITZ, LINDA
/ TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
/ FILE REFERENCE: 700953/47113C
/ CURRENT APPLICATION NUMBER: US/10/057,136
/ CURRENT FILING DATE: 2002-01-25
/ PRIOR APPLICATION NUMBER: 09/366,670
/ PRIOR FILING DATE: 1999-08-03
/ PRIOR APPLICATION NUMBER: PCT/US98/03693
/ PRIOR FILING DATE: 1998-02-24
/ PRIOR APPLICATION NUMBER: 60/038,253
/ PRIOR FILING DATE: 1997-02-24
/ NUMBER OF SEQ ID NOS: 20
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO: 11
/ LENGTH: 60
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-057-136-11

Query Match 48.3%; Score 29; DB 14; Length 60;
Best Local Similarity 71.7%; Pred. No. 0.45;
Matches 38; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 7 ACCGCTCCACCTGCACAGCGGGGTCAACAGCGCCGACACTGACCTGGCC 59
DB 7 ACGGCACTCCACAGCACCGAGTCACTGTGACCCGACACCCGCTCAGCTCC 59

Search completed: January 16, 2005, 09:30:30
Job time : 188.4 secs

This Page Blank (uspto)

Matches	40;	Conservative	0;	Mismatches	17;	Indels	0;	Gaps	0;
Oy	1	GGAGTATCCGCTTCACCTGCACACGGGGTCAACAGCGCGACAGACTGACCTGGC	57						
Db	24	GAGGGGTCCCTCTATCTCTCTCCTCACCATGTCACTAGCGTTGCACACCCGACCTGTG	80						
RESULT 2									
BY729679/c									
LOCUS									
DEFINITION	BY729679			657 bp	mRNA	linear	EST 17-DEC-2002		
ACCESSION	BY729679								
VERSION	BY729679								
KEYWORDS	EST.								
SOURCE	BY729679.1			GI:27142806					
ORGANISM	Mus musculus (house mouse)								
	Mus musculus								

RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer. *Genome Res*
 10 (11): 1157-1771 (2000)
 Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. *Genome Res.* 11 (2): 281-289 (2001)
 cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.
 Please visit our web site (<http://genome.gsc.riken.go.jp>) for
 further details.
 Location/Qualifiers
 1..657
 source
 /organism="Mus musculus"

TITLE
rogers, U., Birney, E. and Hayashizaki, Y.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
JOURNAL
MEDLINE
PUBMED
22354683
12466851
COMMENT
Contact: Yoshihide Hayashizaki

COMMENT

Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suenhiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@sc.riken.jp, URL: <http://genome.gsc.riken.jp/>
Adachi, J., Aizawa, K., Akimura, T., Arawaka, T., Caminici, P.,
Fukuda, S., Hasehizume, W., Hayashida, K., Hirozane, T., Hori, F.,
Imecani, K., Ishii, Y., Itoh, M., Kageawa, T., Kawai, J., Kojima, Y.,
Kondo, S., Kono, H., Koya, S., Miyazaki, A., Murata, M., Nakanura, M.,
Nomura, K., Nunazaki, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N.,
Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M.,
Takeda, Y., Waki, K., Watanuki, A., Muramatsu, M. and Hayashizaki, Y.
Direct Submission

Computational Analysis of Full-length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1611-1630 (2000)

ORIGIN	<i>gtaaccctcagc acatt</i>	alter bulk excision from Lambda FLC I.
Query Match	47.7%	Score 28.6; DB 6; Length 657;
Best Local Similarity	67.8%	Pred. No. 48;
Matches	40; Conservative	0; Mismatches 19; Indels 0; Gaps 0;
OY	1 GGAAGTACCGCTCCACCTGTCAACAGCGCGCCAGACACTGACCTGGCC	59
Db	485 GAAGGCCCGCGCGCGCCCAAGCGGGGCAAGGCGCGGCCCAACACCCGGGCAAGCGGC	427

RESULT 3					
CE582697/c					
LOCUS					
DEFINITION	CE582697	513 bp	DNA	linear	GSS 28-SEP-2003
ACCESSION	tigr-gss-dog-17000366290475		Dog Library	Canis familiaris genomic,	
VERSION	CE582697				
KEYWORDS	CE582697.1	GI:36899478			
SOURCE	GSS.				
ORGANISM	Canis familiaris (dog)				
	Canis familiaris				
	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.				
REFERENCE	1 Daaes I to 513				
AUTHORS	Kirkness,E.F., Batina,V., Halpern,A.L., Levy,S., Remington,K., Rusch,D.B., Delcher,A.L., Pop,M., Wang,W., Fraser,C.M. and Venter,J.C.				
TITLE	The dog genome: survey sequencing and comparative analysis				
JOURNAL	Science 301 (5641), 1898-1903 (2003)				
MEDLINE	22875432				
PUBMED	14512627				
COMMENT	Contact: Kirkness EF The Institute for Genomic Research Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive				

Rockville, MD 20850, USA
Tel: 301-838-0200
Fax: 301-838-0208
Email: ekirknes@tigr.org

Classes: shotgun.
Location/Qualifiers

FEATURES
source 1..513

/organism="Canis familiaris"
/mol_type="genomic DNA"
/strain="Standard Poodle"
/db_xref="taxon:9615"
/clone_lib="Dog Library"
/note="Site 1: Bexxi; Libraries were prepared from peripheral blood"

ORIGIN

Query Match 47.0%; Score 28.2; DB 9; Length 513;
Best Local Similarity 73.5%; Pred. No. 65;
Matches 36; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 8 CCGCTCCACCTGCACACGGGGTCAACAGCGCCGACACTGCACCTGC 56
|||
424 CCGCGACCTGCATCGGGGTCCCGAGAGTCCAGAGCGCCGACCTGC 376

RESULT 4
CE347100/c 514 bp DNA 1linear GSS 26-SRP-2003

LOCUS tigr-gss-dog-1700033414252 Dog Library Canis familiaris genomic,
DEFINITION genomic survey sequence.
ACCESSION CE347100
VERSION CE347100.1 GI:36177454

KEYWORDS GSS.
SOURCE

ORGANISM Canis familiaris (dog)

Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.

REFERENCE
1 (bases 1 to 514)
Kirknes, E.F., Bafna, V., Halpern, A.L., Levy, S., Remington, K.,
Rusch, D.B., Delcher, A.L., Pop, M., Wang, W., Fraser, C.M. and
Venter, J.C.

The dog genome: survey sequencing and comparative analysis
Science 301 (5641), 1898-1903 (2003)

JOURNAL
MEDLINE
PUBMED
22875432
14512627

COMMENT
Contact: Kirknes EF

The Institute for Genomic Research
Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
Rockville, MD 20850, USA

Tel: 301-838-0200
Fax: 301-838-0208
Email: ekirknes@tigr.org

Classes: shotgun.
Location/Qualifiers

FEATURES
source 1..514

/organism="Canis familiaris"
/mol_type="genomic DNA"
/strain="Standard Poodle"
/db_xref="taxon:9615"
/clone_lib="Dog Library"
/note="Site 1: Bexxi; Libraries were prepared from peripheral blood"

ORIGIN

Query Match 47.0%; Score 28.2; DB 9; Length 514;
Best Local Similarity 73.5%; Pred. No. 65;
Matches 36; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 8 CCGCTCCACCTGCACACGGGGTCAACAGCGCCGACACTGCACCTGC 56
|||
51 CCGCGACCTGCATCGGGGTCCCGAGAGTCCAGAGCGCCGACCTGC 3

RESULT 5
CE362548 603 bp DNA 1linear GSS 27-SRP-2003

LOCUS tigr-gss-dog-17000361549348 Dog Library Canis familiaris genomic,
DEFINITION genomic survey sequence.
ACCESSION CE362548
VERSION CE362548.1 GI:36577281

KEYWORDS GSS.
SOURCE

ORGANISM Canis familiaris (dog)

Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.

REFERENCE
1 (bases 1 to 603)
Kirknes, E.F., Bafna, V., Halpern, A.L., Levy, S., Remington, K.,
Rusch, D.B., Delcher, A.L., Pop, M., Wang, W., Fraser, C.M. and
Venter, J.C.

The dog genome: survey sequencing and comparative analysis
Science 301 (5641), 1898-1903 (2003)

JOURNAL
MEDLINE
PUBMED
22875432
14512627

COMMENT
Contact: Kirknes EF

The Institute for Genomic Research
Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
Rockville, MD 20850, USA

Tel: 301-838-0200
Fax: 301-838-0208
Email: ekirknes@tigr.org

Classes: shotgun.
Location/Qualifiers

FEATURES
source 1..603

/organism="Canis familiaris"
/mol_type="genomic DNA"
/strain="Standard Poodle"
/db_xref="taxon:9615"
/clone_lib="Dog Library"
/note="Site 1: Bexxi; Libraries were prepared from peripheral blood"

ORIGIN

Query Match 47.0%; Score 28.2; DB 9; Length 603;
Best Local Similarity 73.5%; Pred. No. 65;
Matches 36; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 8 CCGCTCCACCTGCACACGGGGTCAACAGCGCCGACACTGCACCTGC 56
|||
141 CCGCGACCTGCATCGGGGTCCCGAGAGTCCAGAGCGCCGACCTGC 189

RESULT 6
BX611440 652 bp mRNA 1linear EST 07-AUG-2003

LOCUS BX611440 Normalized Anopheles Head (NAH) Library Anopheles gambiae
DEFINITION cDNA clone AGAA781TR, mRNA sequence.
ACCESSION BX611440
VERSION BX611440.1 GI:33501327

KEYWORDS EST.
SOURCE

ORGANISM Anopheles gambiae (African malaria mosquito)

Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;
Anopheles.

REFERENCE
1 (bases 1 to 652)
Lobo, N.L., Gardner, M., Romans, P. and Collins, F.H.

Anopheles gambiae EST, Center for Tropical Disease Research and
Training
Unpublished (2003)

Center for Tropical Disease Research and Training
University of Notre Dame
Notre Dame, IN 46556, USA

Tel: 574-631-9245
Fax: 574-631-3996
Email: frank.h.collins.75@nd.edu.

FEATURES

source

Location/Qualifiers

1..652
/organism="Anopheles gambiae"
/mol_type="mRNA"
/db_xref="taxon:7165"
/clone="AGAA7817R"
/lab_host="E. coli DH10B"
/note="Vector: p773D-Pac (Pharmacia) with a modified directional site 1: EcoRI (5' end); Site 2: NotI (3' end); a library constructed from strain 4art adult mosquito heads. Equal numbers of sugar fed males, sugar fed females and 6, 24 and 48 hr post blood meal females were used: Bonaldo, Lennon & Soares (1996). Normalization and Subtraction: Two Approaches to Facilitate Gene Discovery, Genome Research 6: 791-806. ESTs sequenced from the M13 reverse priming site reading from the 5' ends of the cDNAs are indicated by 'R' in the clone name. ESTs sequenced from the M13 forward priming site reading from the 3' ends of the cDNAs are indicated by 'F' in the clone name."

ORIGIN

Query Match 47.0%; Score 28.2; DB 5; Length 652;
Best Local Similarity 68.4%; Pred. No. 66;
Matches 39; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1 GGAACTACCGCTTCACCTGACACGCGGGTCAAGCGCGGACCTCGACCTGGC 57
DB 395 GGAGGGTCCCTCTATCTGCTCACCACATCTACCTGACGCGCCGACCTGTG 451

RESULT 7
LOCUS BM607934 682 bp mRNA linear EST 25-FEB-2002
DEFINITION 17000687085787 A.Gam.ad.cDNA.blood1 Anopheles gambiae cDNA clone
ACCESSION 1960049719648 5, mRNA sequence.
VERSION BM607934
KEYWORDS BM607934.1 GI:18906038
SOURCE EST.
ORGANISM Anopheles gambiae (African malaria mosquito)
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;
Anopheles.

REFERENCE 1 (bases 1 to 682)
Holt, R.A., Lin, J.-J., Murphy, S.D., Evans, C.A., Kraft, C.L., Charlab, R., Collins, F.H., Venter, D.C. and Hoffman, S.L.,
Celera Anopheles gambiae EST project
Unpublished (2002)
Contact: Holt R.A.
Celera Genomics
45 W. Gude Dr., Rockville, MD 20850, USA
Tel: 2404531151
Fax: 2404534580
Email: HoltR@celera.com
Plate: NU010049VC row: L column: 02
Seq primer: M13 Reverse.

FEATURES

source

Location/Qualifiers

1..682
/organism="Anopheles gambiae"
/mol_type="mRNA"
/strain="RSP-ST (Reduced suscep. to Permethrin - std. chromosome)"
/db_xref="taxon:7165"
/clone="19600449719648"
/dev_stage="Adult"
/lab_host="DH10B"
/clone_lib="A.Gam.ad.cDNA.blood1"
/note="Vector: pSport1, Site 1: SalI; Site 2: NotI; Whole adult mosquitoes (mixed sex) frozen on liquid nitrogen 24 hours after human blood feeding. cDNA inserts >500 bp cloned directionally into pSport 1. Not 1 site is 3'.

ORIGIN

Query Match 47.0%; Score 28.2; DB 4; Length 682;
Best Local Similarity 68.4%; Pred. No. 66;
Matches 39; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1 GGAACTACCGCTTCACCTGACACGCGGGTCAAGCGCGGACCTCGACCTGGC 57
DB 336 GGAGGGTCCCTCTATCTGCTCACCACATCTACCTGACGCGCCGACCTGTG 392

RESULT 8
LOCUS BG005521/c 583 bp mRNA linear EST 24-JAN-2001
DEFINITION WR3-GN0186-241100-008-b11 GN0186 Homo sapiens cDNA, mRNA sequence.
ACCESSION BG005521
VERSION BG005521.1 GI:12447774
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 583)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H., Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.G.
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
10737800
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rita Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the PAPSP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR3&cl=MR3-GN0186-241100-008-b11&cl3=2000-11-24&cl4=1)
Seq primer: puc 18 forward
High quality sequence stop: 539.

FEATURES

source

Location/Qualifiers

1..583
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="GN0186"
/note="Organ: placenta, normal; Vector: puc18, Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORFESTS PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Query Match 46.3%; Score 27.8; DB 4; Length 583;
Best Local Similarity 74.5%; Pred. No. 88;
Matches 35; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 13 CCACCTGACACGCGGGTCAAGCGCGGACCTCGACCTGGC 59
DB 83 CCACCTGACACATTCCTCCACACGCGGACCTCGACACGCGCC 37

RESULT 9
 A1925867/c 330 bp mRNA linear EST 08-MAR-2000
 LOCUS wo20d04.x1 NCI CGAP Pauli Homo sapiens cDNA clone IMAGE:2455879 3'
 DEFINITION similar to contains element MSRI MSRI repetitive element ;, mRNA
 sequence.
 ACCESSION A1925867
 VERSION A1925867.1 GI:5661831
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 330)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Life Technologies catalog #: 11548-013
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bdrp/image/image.html
 Insert Length: 578 Std Error: 0.00
 Seq primer: -40UP from gibco
 High quality sequence stop: 305.
 Location/Qualifiers
 1..330
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:2455879"
 /tissue_type="adenocarcinoma"
 /lab_host="DH10B"
 /clone_lib="NCI-CGAP_Pauli"
 /note="Organ: pancreas; Vector: pCMV-SPORT6; Site_1: SalI;
 Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
 Average insert size 1.72 kb. Life Technologies catalog #:
 11548-013"

ORIGIN

Query Match 45.7%; Score 27.4; DB 1; Length 330;
 Best Local Similarity 69.8%; Pred. No. 1.2e+02;
 Matches 37; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 7 ACCGCTCCACCTGCACACGGGGTCAACAGCGCCGACACCTGCGCC 59
 |||||
 DB 310 ACCGCCCCCAGCCGCGGTGTCTGCGCGCCCGGACACGCGCGCC 258

RESULT 10
 LOCUS BM791359 604 bp mRNA linear EST 05-MAR-2002
 DEFINITION K-EST0071342 S21SNUS20 Homo sapiens cDNA clone S21SNUS20-14-A06 5'
 mRNA sequence.
 ACCESSION BM791359
 VERSION BM791359.1 GI:19139591
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 604)
 Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
 Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
 Kim,Y.S.
 21C Frontier Korean EST Project 2001
 JOURNAL Unpublished (2002)
 COMMENT Contact: Kim YS

Genome Research Center
 Korea Research Institute of Bioscience & Biotechnology
 52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
 Tel: +82-42-860-4470
 Fax: +82-42-860-4409
 Email: yongsung@mail.kribb.re.kr
 Plate: 14 row: A column: 06
 High quality sequence stop: 604.
 Location/Qualifiers
 1..604
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="S21SNUS20-14-A06"
 /sex="F"
 /tissue_type="Stomach"
 /cell_type="Floating aggregates"
 /cell_line="SNU-520"
 /lab_host="Top10F"
 /clone_lib="S21SNUS20"
 /note="Organ: Stomach; Vector: pTZ18RP1; Site_1: EcoRI;
 Site_2: NotI; The poly (A) + RNA was dephosphorylated with
 bacterial alkaline phosphatase (BAP) and then deacapped
 with tabacco acid pyrophosphatase (TAP). The deacapped
 intact mRNA was ligated with DNA-RNA linker including EcoR
 I site by treatment of T4 RNA ligase and the first strand
 cDNA was synthesized from oligo dt-selected mRNA by
 priming with dt-tailed vector. The dt-tailed vector was
 adjusted to have about 60nt. The cDNA vector was
 circularized with E. coli DNA ligase after digestion of
 EcoRI which site is also included in vector. An RNA strand
 converted to a DNA strand by Okayama-Berg method. The
 obtained cDNA vectors were used for transformation of
 competent cells E. coli Top10F by electroporation method.
 The cDNA libraries constructed by this method are
 full-length enriched cDNA library."

ORIGIN

Query Match 45.7%; Score 27.4; DB 4; Length 604;
 Best Local Similarity 69.8%; Pred. No. 1.2e+02;
 Matches 37; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 7 ACCGCTCCACCTGCACACGGGGTCAACAGCGCCGACACCTGCGCC 59
 |||||
 DB 34 ACCGACCCCGCAGCCGCGGTGTCTGCGCGCCCGGACACGCGCC 86

RESULT 11
 LOCUS BUS42454 877 bp mRNA linear EST 13-SEP-2002
 DEFINITION AGENCOURT_10322173 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6574322
 5', mRNA sequence.
 ACCESSION BUS42454
 VERSION BUS42454.1 GI:22852937
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 877)
 NIH-MGC <http://mgs.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Contact: Robert Strausberg, Ph.D.
 Tissue Procurement: DCTD/DRP
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: L1CM2769 row: h column: 02

FEATURES High quality sequence stop: 760.
Location/Qualifiers
source 1..877

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6574323"
/tissue_type="carcinoma, cell line"
/lab_host="MD10B (phage-resistant)"
/clone_id="NIH_MGC_40"
/note="Organ: prostate; Vector: pOT7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 45.7%; Score 27.4; DB 5; Length 877;
Best Local Similarity 69.8%; Pred. No. 1.2e+02;
Matches 37; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 7 ACCGCTCCAGCTGCACAGGGGTCAAGCGCCGACAGACTCGACCTGCGCC 59
Db 41 ACCGCCCCCAGCCAGGTGTCACTCGGCCCCGACACAGCGCCGCCCC 93

RESULT 12 CB994654 890 bp mRNA linear EST 01-MAY-2003
LOCUS IMAGE:30335618 5', mRNA sequence.
DEFINITION CB994654
ACCESSION CB994654
VERSION CB994654.1 GI:30289174
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.
REFERENCE 1 (bases 1 to 890)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Dr. Stefan Hansson
CDNA Library Preparation: Michael J. Brownstein (NIH/NIHRI) with help and advice from Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAM357 row: m column: 03
High quality sequence stop: 503.
Location/Qualifiers

FEATURES

source

1..890
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30335618"
/tissue_type="pre-eclampic placenta"
/lab_host="MD10B Tona"
/clone_id="NIH_MGC_148"
/note="Organ: placenta; Vector: pBluescriptR; Site 1: all-XhoI; Site 2: BamHI; Library is oligo-dT primed and directionally cloned using primer 5'-TTTTTTTTTTTTTNN-3', size-selected for average insert size 2.3 kb and normalized to ROP 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation)."

ORIGIN

Library constructed by M. Brownstein (NIH/NIHRI).
National Institutes of Health). Note: this is a NIH_MGC Library."

Query Match 45.7%; Score 27.4; DB 6; Length 890;
Best Local Similarity 75.6%; Pred. No. 1.2e+02;
Matches 34; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 GGAGTACCGCTCCACCTGCACAGGGGTCAAGCGCCGACAC 45
Db 721 GGTGTACCGATCACTCCCGAGAGTGTCAAGGCGCCGAC 765

RESULT 13 CA489836 959 bp mRNA linear EST 14-NOV-2002
LOCUS IMAGE:6722324 5',
DEFINITION CA489836
ACCESSION CA489836
VERSION CA489836.1 GI:24952627
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.
REFERENCE 1 (bases 1 to 959)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Kristi A. Eglund, Ira Pastan
CDNA Library Preparation: Invitrogen Corp
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM4284 row: n column: 20
High quality sequence start: 31
High quality sequence stop: 446.
Location/Qualifiers

FEATURES

source

1..959
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6722324"
/cell_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231, hTERT-HM1, LNCaP"
/lab_host="EMD10B"
/clone_id="MABCL"
/note="Vector: pCMV-SPORT6; Site 1: EcoRV; Site 2: Not I; Subtracted with brain, liver, lung, kidney and muscle. Directionally cloned. Priming method: oligo-dT. Average insert size: 1800 bp. Library amplification: 26,000 fold. Kristi A. Eglund, James J. Vincent, Robert Strausberg, Bungkok Lee & Ira Pastan. Discovery of new breast cancer genes encoding membrane and secreted proteins. Manuscript submitted."

ORIGIN

Query Match 45.7%; Score 27.4; DB 6; Length 959;
Best Local Similarity 69.8%; Pred. No. 1.2e+02;
Matches 37; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 7 ACCGCTCCAGCTGCACAGGGGTCAAGCGCCGACAGACTCGACCTGCGCC 59
Db 355 ACCGCCCCCAGCCAGGTGTCACTCGGCCCCGACACAGCGCCGCCCC 407

RESULT 14 BU148487

LOCUS BU148487 1113 bp mRNA linear EST 03-SEP-2002
 DEFINITION AGENCOURT 8670479 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6380642
 5', mRNA Sequence.
 ACCESSION BU148487
 VERSION BU148487.1 GI:22662019
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 1113)
 NIH-MGC http://mgi.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: DCTD/DRP
 cDNA Library Preparation: Rubin Laboratory
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L1CM2569 row: 1 column: 03
 High quality sequence stop: 235.
 Location/Qualifiers
 1..1113
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6380642"
 /tissue_type="carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_11b="NIH_MGC_40"
 /note="Organ: prostate; Vector: pOTB7; Site: 1: XhoI;
 Site: 2: EcoRI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCACGAG(G). Library constructed by
 Ling Hong in the laboratory of Gerald M. Rubin (University
 of California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library."

ORIGIN

Query Match 45.7%; Score 27.4; DB 5; Length 1113;
 Best Local Similarity 69.8%; Pred. No. 1.2e+02;
 Matches 37; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 7 ACCGCTCCACCTGCACAGCGGTCACAGCGGCCGACACCTCGCGCC 59
 |||||
 DB 36 ACCGCCCCCCCAGCCACGCTGTCACTCTGCCCCGACACCAAGCGGCC 88
 |||||

RESULT 15
 BUS42996 1130 bp mRNA linear EST 13-SEP-2002
 LOCUS AGENCOURT_10338707 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6574979
 DEFINITION 5', mRNA Sequence.
 ACCESSION BUS42996
 VERSION BUS42996.1 GI:22853479
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 1130)
 NIH-MGC http://mgi.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: DCTD/DRP
 cDNA Library Preparation: Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L1CM2771 row: c column: 11
 High quality sequence start: 27
 High quality sequence stop: 246.
 Location/Qualifiers
 1..1130
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6574979"
 /tissue_type="carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_11b="NIH_MGC_40"
 /note="Organ: prostate; Vector: pOTB7; Site: 1: XhoI;
 Site: 2: EcoRI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCACGAG(G). Library constructed by
 Ling Hong in the laboratory of Gerald M. Rubin (University
 of California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library."

ORIGIN

Query Match 45.7%; Score 27.4; DB 5; Length 1130;
 Best Local Similarity 69.8%; Pred. No. 1.2e+02;
 Matches 37; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 7 ACCGCTCCACCTGCACAGCGGTCACAGCGGCCGACACCTCGCGCC 59
 |||||
 DB 75 ACCGCCCCCCCAGCCACGCTGTCACTCTGCCCCGACACCAAGCGGCC 127
 |||||

Search completed: January 16, 2005, 02:55:36
 Job time : 1553.6 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2005, 19:50:13 : Search time 35.2 Seconds
(without alignments)
1211.572 Million cell updates/sec

Title: US-10-057-136-9

Perfect score: 60
Sequence: 1 ggaagtagcgcctccacatgcg.....cagacactgcactgcgcca 60

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/prodata/1/ina/5A_COMB.seq: *
2: /cgn2_6/prodata/1/ina/5B_COMB.seq: *
3: /cgn2_6/prodata/1/ina/6A_COMB.seq: *
4: /cgn2_6/prodata/1/ina/6B_COMB.seq: *
5: /cgn2_6/prodata/1/ina/PCTUS_COMB.seq: *
6: /cgn2_6/prodata/1/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the distribution.
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	27.4	45.7	1804	US-10-029-517-17	Sequence 17, App1
2	27.4	45.7	8186	US-10-029-517-19	Sequence 19, App1
3	26	43.3	518	US-10-029-517-101	Sequence 101, App
4	25.8	43.0	572	US-10-029-517-18	Sequence 18, App1
5	25.8	43.0	1721	US-10-029-517-3	Sequence 3, App1
6	24.6	41.0	60	US-09-475-947A-246	Sequence 246, App
7	24.4	40.7	846	US-09-455-960-5	Sequence 5, App1
8	24.4	40.7	846	US-10-051-325-5	Sequence 5, App1
9	24.4	40.7	2500	US-09-455-960-2	Sequence 2, App1
10	24.4	40.7	3343	US-10-051-325-2	Sequence 102, App
11	23.6	39.3	1392	US-09-489-039A-5663	Sequence 86, App
12	23.6	39.3	1772	US-09-276-531-86	Sequence 86, App
13	23.2	38.7	507	US-08-956-171E-3895	Sequence 3895, App
14	23.2	38.7	507	US-08-781-986A-3895	Sequence 3895, App
15	23.2	38.7	558	US-08-259-000-1	Sequence 1, App1
16	23.2	38.7	558	US-08-729-767-1	Sequence 1, App1
17	23.2	38.7	657	US-09-252-991A-15827	Sequence 15827, A
18	23.2	38.7	3280	US-10-037-616-2	Sequence 2, App1
19	23.2	38.7	3280	US-08-259-000-4	Sequence 2, App1
20	23.2	38.7	3280	US-08-729-767-6	Sequence 6, App1
21	23.2	38.7	3583	US-09-252-991A-15893	Sequence 15893, A
22	23.2	38.7	4257	US-09-252-991A-15796	Sequence 15796, A
23	23.2	38.7	6192	US-08-479-537A-1	Sequence 1, App1
24	23.2	38.7	6192	US-09-083-116-1	Sequence 1, App1
25	23.2	38.7	6192	US-09-134-916A-1	Sequence 1, App1
26	23.2	38.7	6449	US-08-479-537A-4	Sequence 4, App1
27	23.2	38.7	6449	US-08-479-537A-4	Sequence 4, App1

28	23.2	38.7	6449	US-09-083-116-4	Sequence 4, App1
29	23.2	38.7	6449	US-09-134-916A-4	Sequence 4, App1
30	23.2	38.7	13794	US-08-956-171E-54	Sequence 54, App1
31	23.2	38.7	13794	US-08-781-986A-54	Sequence 54, App1
32	23	38.3	461	US-09-621-976-3353	Sequence 3353, App
33	23	38.3	2680	US-09-566-921-9	Sequence 9, App1
34	23	38.3	2727	US-09-919-497-13	Sequence 13, App1
35	22.8	38.0	425	US-08-956-171E-3693	Sequence 3693, App
36	22.8	38.0	425	US-08-781-986A-3693	Sequence 3693, App
37	22.8	38.0	468	US-09-252-991A-682	Sequence 682, App
38	22.8	38.0	780	US-09-252-991A-729	Sequence 729, App
39	22.8	38.0	792	US-09-489-039A-1068	Sequence 1068, App
40	22.8	38.0	981	US-10-029-517-16	Sequence 16, App1
41	22.8	38.0	1221	US-09-461-325-471	Sequence 471, App
42	22.8	38.0	1221	US-10-012-542-471	Sequence 471, App
43	22.8	38.0	1221	US-10-115-123-471	Sequence 471, App
44	22.8	38.0	1379	US-09-461-325-124	Sequence 124, App
45	22.8	38.0	1379	US-10-012-542-124	Sequence 124, App

ALIGNMENTS

```
RESULT 1
US-10-029-517-17
; Sequence 17, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Doble
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 17
; LENGTH: 1804
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (73) ... (1500)
US-10-029-517-17

Query Match          45.7%; Score 27.4; DB 4; Length 1804;
Best Local Similarity 69.8%; Pred. No. 0.85;
Matches 37; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY      7 ACCGCTCCACCTGCACACGGGCTCCACACGGGCGGCGACACTCGACCTGGCC 59
Db      463 ACCGCCCTCCACGACCTGCTCCTCGGCGGCGACACGCGGCGGCGGCGG 515

RESULT 2
US-10-029-517-19
; Sequence 19, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Doble
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 19
; LENGTH: 8186
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 6899
; OTHER INFORMATION: unknown
```

```
/ NAME/KEY: unsure
/ LOCATION: 7155
/ OTHER INFORMATION: unknown
/ NAME/KEY: unsure
/ LOCATION: 7184
/ OTHER INFORMATION: unknown
/ NAME/KEY: unsure
/ LOCATION: 7957
/ OTHER INFORMATION: unknown
/ NAME/KEY: inttron
/ LOCATION: (2997)...(3498)
/ OTHER INFORMATION: inttron 1
/ NAME/KEY: inttron:exon junction
/ LOCATION: (3498)...(3499)
/ OTHER INFORMATION: inttron 1:exon 2
/ NAME/KEY: exon
/ LOCATION: (3508)...(3599)
/ OTHER INFORMATION: exon 2d
/ NAME/KEY: exon:inttron junction
/ LOCATION: (3982)...(3983)
/ OTHER INFORMATION: exon 2a:inttron 2a
/ NAME/KEY: inttron:exon junction
/ LOCATION: (4205)...(4206)
/ OTHER INFORMATION: inttron 2c:exon 3c
/ NAME/KEY: inttron:exon junction
/ LOCATION: (4259)...(4260)
/ OTHER INFORMATION: inttron 2d:exon 3d
/ NAME/KEY: exon
/ LOCATION: (4260)...(4328)
/ OTHER INFORMATION: exon 3d
/ NAME/KEY: inttron:exon junction
/ LOCATION: (4632)...(4633)
/ OTHER INFORMATION: inttron 3:exon 4
/ NAME/KEY: exon
/ LOCATION: (4914)...(5035)
/ OTHER INFORMATION: exon 5
/ NAME/KEY: inttron
/ LOCATION: (5266)...(6293)
/ OTHER INFORMATION: inttron 6
US-10-029-517-19
```

```
Query Match 45.7%; Score 27.4; DB 4; Length 8186;
Best Local Similarity 69.8%; Pred. No. 1.1;
Matches 37; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
```

```
QY 7 ACCGCTCCACCTGCACACGGGGTCAACAAGCGCCGACACTGCAGCTGGGCC 59
DB 3831 ACCGCCCCCCCCACGCGGTGTCACTCGGCCCCCGGACACCAAGCGGCCCC 3883
```

RESULT 3

```
US-10-029-517-101
/ Sequence 101, Application US/10029517
/ Patent No. 6716627
/ GENERAL INFORMATION:
/ APPLICANT: Kenneth W. Dobie
/ TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
/ FILE REFERENCE: RTS-0352
/ CURRENT APPLICATION NUMBER: US/10/029,517
/ NUMBER OF SEQ ID NOS: 107
/ SEQ ID NO 101
/ LENGTH: 518
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
US-10-029-517-101
```

```
Query Match 43.3%; Score 26; DB 4; Length 518;
Best Local Similarity 70.0%; Pred. No. 2.2;
Matches 35; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
```

```
QY 7 ACCGCTCCACCTGCACACGGGGTCAACAAGCGCCGACACTGCAGCTGGC 56
DB 468 ACCGCCCCCCCCACGCGGTGTCACTCGGCCCCCGGACACCAAGCGGCC 517
```

RESULT 4

```
US-10-029-517-18
/ Sequence 18, Application US/10029517
/ Patent No. 6716627
/ GENERAL INFORMATION:
/ APPLICANT: Kenneth W. Dobie
/ TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
/ FILE REFERENCE: RTS-0352
/ CURRENT APPLICATION NUMBER: US/10/029,517
/ NUMBER OF SEQ ID NOS: 107
/ SEQ ID NO 18
/ LENGTH: 572
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (67)...(572)
US-10-029-517-18
```

```
Query Match 43.0%; Score 25.8; DB 4; Length 572;
Best Local Similarity 67.9%; Pred. No. 2.6;
Matches 36; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
```

```
QY 7 ACCGCTCCACCTGCACACGGGGTCAACAAGCGCCGACACTGCAGCTGGGCC 59
DB 484 ACCGCCCCCCCCACGCGGTGTCACTCGGCCCCCGGACACCAAGCGGCC 536
```

RESULT 5

```
US-10-029-517-3
/ Sequence 3, Application US/10029517
/ Patent No. 6716627
/ GENERAL INFORMATION:
/ APPLICANT: Kenneth W. Dobie
/ TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
/ FILE REFERENCE: RTS-0352
/ CURRENT APPLICATION NUMBER: US/10/029,517
/ NUMBER OF SEQ ID NOS: 107
/ SEQ ID NO 3
/ LENGTH: 1721
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (58)...(1605)
US-10-029-517-3
```

```
Query Match 43.0%; Score 25.8; DB 4; Length 1721;
Best Local Similarity 67.9%; Pred. No. 3.1;
Matches 36; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
```

```
QY 7 ACCGCTCCACCTGCACACGGGGTCAACAAGCGCCGACACTGCAGCTGGGCC 59
DB 448 ACCGCCCCCCCCACGCGGTGTCACTCGGCCCCCGGACACCAAGCGGCC 500
```

RESULT 6

```
US-09-475-947A-246
/ Sequence 246, Application US/09475947A
/ Patent No. 6472154
/ GENERAL INFORMATION:
/ APPLICANT: Garner, Harold R.
/ APPLICANT: Wren, Jonathan D.
/ APPLICANT: Minna, John D.
```



```

; TITLE OF INVENTION: Polymorphic Repeats in Human Genes
; FILE REFERENCE: UTS0667
; CURRENT APPLICATION NUMBER: US/09/475,947A
; CURRENT FILING DATE: 1999-12-31
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 246
; LENGTH: 60
; TYPE: DNA
; ORGANISM: human
US-09-475-947A-246

Query Match          41.0%; Score 24.6; DB 4; Length 60;
Best Local Similarity 70.2%; Pred. No. 4.7;
Matches 33; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy      7  ACCGCTCAGCTGACACGGGGTCAACAGCGCGCCAGACACTGACG 53
Db      13  ACCGCCCCCGACAGCCGAGTGTCACTCGGCCCGGACACCAAGGCC 59

RESULT 7
US-09-455-960-5/c
; Sequence 5, Application US/09455960
; Patent No. 6361776
; GENERAL INFORMATION:
; APPLICANT: Delcayre, Alain
; TITLE OF INVENTION: Compositions isolated from M. vaccae and
; FILE REFERENCE: 11000.1047
; CURRENT APPLICATION NUMBER: US/09/455,960
; CURRENT FILING DATE: 1999-12-06
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 846
; TYPE: DNA
; ORGANISM: Mycobacterium vaccae
US-09-455-960-5

Query Match          40.7%; Score 24.4; DB 3; Length 846;
Best Local Similarity 63.8%; Pred. No. 8.7;
Matches 37; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

Qy      1  GGAAGTACCGCTTCACCTGACACAGCGGGTCAACAGCGCCGACACTGACCTGCGC 58
Db      487  GGACGACCGACCGACCGCGGACCCCGGTCAACCGCGCGGAAGATGTGACACGCGC 430

RESULT 8
US-10-051-325-5/c
; Sequence 5, Application US/10051325
; Patent No. 6716430
; GENERAL INFORMATION:
; APPLICANT: Delcayre, Alain
; TITLE OF INVENTION: Compositions isolated from M. vaccae and
; FILE REFERENCE: 11000.1047c1
; CURRENT APPLICATION NUMBER: US/10/051,325
; CURRENT FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US 09/455,960
; PRIOR FILING DATE: 1999-12-06
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 846
; TYPE: DNA
; ORGANISM: Mycobacterium vaccae
US-10-051-325-5

Query Match          40.7%; Score 24.4; DB 4; Length 846;
Best Local Similarity 63.8%; Pred. No. 8.7;
Matches 37; Conservative 0; Mismatches 21; Indels 0; Gaps 0;
```

```

Qy      1  GGAAGTACCGCTTCACCTGACACAGCGGGTCAACAGCGCGCCAGACACTGACCTGCGC 58
Db      487  GGACGACCGACCGACCGCGGACCCCGGTCAACCGCGCGGAAGATGTGACACGCGC 430

RESULT 9
US-09-455-960-2/c
; Sequence 2, Application US/09455960
; Patent No. 6361776
; GENERAL INFORMATION:
; APPLICANT: Delcayre, Alain
; TITLE OF INVENTION: Compositions isolated from M. vaccae and
; FILE REFERENCE: 11000.1047
; CURRENT APPLICATION NUMBER: US/09/455,960
; CURRENT FILING DATE: 1999-12-06
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Mycobacterium vaccae
US-09-455-960-2

Query Match          40.7%; Score 24.4; DB 3; Length 2500;
Best Local Similarity 63.8%; Pred. No. 10;
Matches 37; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

Qy      1  GGAAGTACCGCTTCACCTGACACAGCGGGTCAACAGCGCGCCAGACACTGACCTGCGC 58
Db      2141  GGACGACCGACCGACCGCGGACCCCGGTCAACCGCGCGGAAGATGTGACACGCGC 2084

RESULT 10
US-10-051-325-2/c
; Sequence 2, Application US/10051325
; Patent No. 6716430
; GENERAL INFORMATION:
; APPLICANT: Delcayre, Alain
; TITLE OF INVENTION: Compositions isolated from M. vaccae and
; FILE REFERENCE: 11000.1047c1
; CURRENT APPLICATION NUMBER: US/10/051,325
; CURRENT FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US 09/455,960
; PRIOR FILING DATE: 1999-12-06
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Mycobacterium vaccae
US-10-051-325-2

Query Match          40.7%; Score 24.4; DB 4; Length 2500;
Best Local Similarity 63.8%; Pred. No. 10;
Matches 37; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

Qy      1  GGAAGTACCGCTTCACCTGACACAGCGGGTCAACAGCGCGCCAGACACTGACCTGCGC 58
Db      2141  GGACGACCGACCGACCGCGGACCCCGGTCAACCGCGCGGAAGATGTGACACGCGC 2084

RESULT 11
US-10-029-517-102
; Sequence 102, Application US/10029517
; Patent No. 6716827
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0052
```

;;
;; CURRENT APPLICATION NUMBER: US/10/029,517
;;
;; CURRENT FILING DATE: 2001-12-20
;;
;; NUMBER OF SEQ ID NOS: 107
;;
;; SEQ ID NO: 102
;;
;; LENGTH: 3343
;;
;; TYPE: DNA
;;
;; ORGANISM: Homo sapiens
;;
;; FEATURE:
US-10-029-517-102

Query Match
Best Local Similarity 40.7%; Score 24.4; DB 4; Length 3343;
Matches 34; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Db
1734 ACCGCTCCACCTGCACAGCGGGTCAACAAGCGCCGACACTGCACCTGC 56
1734 ACCGCCCCCGACCGCATGTGTACCTCGGCGCCGACACAGCGCCGC 1783

RESULT 12
US-09-489-039A-5663
Sequence 5663, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:
APPLICANT: Gary Breton et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
FILE REFERENCE: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
CURRENT APPLICATION NUMBER: US/09/489,039A
PRIOR FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
NUMBER OF SEQ ID NOS: 1999-01-29
SEQ ID NO 5663
LENGTH: 1392
TYPE: DNA
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-5663

Query Match
Best Local Similarity 39.3%; Score 23.6; DB 4; Length 1392;
Matches 32; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Db
15 ACCGTGACACGGGGTCAACAAGCGCCGACACTGCACCGCGCCCA 60
980 ACCCGCGCGCGCTGTCCGACAGCGCGCCGCGCTGATGTGCGA 1025

RESULT 13
US-09-276-531-86/c
Sequence 86, Application US/09276531
Patent No. 6183968
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Lal, Preeti
APPLICANT: Hillman, Jennifer L.
APPLICANT: Yue, Henry
APPLICANT: Reddy, Roopa
APPLICANT: Guegler, Karl J.
APPLICANT: Baughn, Mariah R.
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF GENES ENCODING
NUMBER OF SEQUENCES: 134
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

;;
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;;
;; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
;;
;; CURRENT APPLICATION DATA:
;;
;; APPLICATION NUMBER: US/09/276,531
;;
;; FILING DATE: Herewith
;;
;; CLASSIFICATION:
;;
;; PRIOR APPLICATION DATA:
;;
;; APPLICATION NUMBER: 60/079,677
;;
;; FILING DATE: March 27, 1998
;;
;; CLASSIFICATION:
;;
;; ATTORNEY/AGENT INFORMATION:
;;
;; NAME: Lynn E. Murry, Ph.D.
;; REGISTRATION NUMBER: 42,918
;; REFERENCE/DOCKET NUMBER: PA-0008 US
;; TELECOMMUNICATION INFORMATION:
;;
;; TELEPHONE: (650) 855-0555
;;
;; TELEFAX: (650) 845-4166
;;
;; INFORMATION FOR SEQ ID NO: 86:
;;
;; SEQUENCE CHARACTERISTICS:
;;
;; LENGTH: 1772 base pairs
;;
;; TYPE: nucleic acid
;;
;; STRANDEDNESS: single
;;
;; TOPOLOGY: linear
;;
;; IMMEDIATE SOURCE:
;;
;; LIBRARY: PROSTUT09
;;
;; CLONE: 1646005
US-09-276-531-86

Query Match
Best Local Similarity 39.3%; Score 23.6; DB 3; Length 1772;
Matches 35; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Db
627 ACCGCTCCACGATCCGACAGGGGCTCCGACACCGCTGCGACTGCACCTGCAGCA 574

RESULT 14
US-08-956-171E-3895/c
Sequence 3895, Application US/08956171E
Patent No. 6593114
GENERAL INFORMATION:
APPLICANT: Charles Kunesh
Gil H. Choi
Patrick S. Dillon
Craig A. Rosen
Steven C. Barash
Michael R. Fannon
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5256
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/956,171E
FILING DATE: 20-Oct-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/009,861
FILING DATE: January 5, 1996
APPLICATION NUMBER: 08/781,986
FILING DATE: January 3, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Mark J. Hyman

Search completed: January 16, 2005, 03:01:43
Job time : 36.2 secs

REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB248P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (240) 314-1224
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 3895:
SEQUENCE CHARACTERISTICS:
LENGTH: 507 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-956-171E-3895

Query Match 38.7%; Score 23.2; DB 4; Length 507;
Best Local Similarity 61.7%; Pred. No. 21;
Matches 37; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Oy 1 GGAAGTACCGCTCCACCTGACACAGGCGGCAGACACTGACCTGGCCCA 60
Db 279 GAAACACCAACACCGACCGACACCGAAGTGGCGAGCCAGAACTCCACACCGCCA 220

RESULT 15

US-08-781-986A-3895/C
Sequence 3895 Application US/08781986A
Patent No. 6737248
GENERAL INFORMATION:
APPLICANT: Charles Kunesh
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5255
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781,986A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Benson, Bob
REGISTRATION NUMBER: 30,446
REFERENCE/DOCKET NUMBER: PB248PP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 3895:
SEQUENCE CHARACTERISTICS:
LENGTH: 507 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-781-986A-3895

Query Match 38.7%; Score 23.2; DB 4; Length 507;
Best Local Similarity 61.7%; Pred. No. 21;
Matches 37; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Oy 1 GGAAGTACCGCTCCACCTGACACAGGCGGCAGACACTGACCTGGCCCA 60
Db 279 GAAACACCAACACCGACCGACACCGAAGTGGCGAGCCAGAACTCCACACCGCCA 220

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2005, 16:00:11 : Search time 170.5 Seconds
(without alignments)
1847.304 Million cell updates/sec

Title: US-10-057-136-9

Perfect score: 60

Sequence: 1 ggaagtagcgcctccacctgc.....cagacacccgacctgcgcca 60

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_23Sep04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001s:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	60	2	AAV48322 Nucleotid
2	52	86.7	60	2	AAV48321 Nucleotid
3	38.2	63.7	60	2	AAV48325 Nucleotid
4	36	60.0	1371	3	AAV48328 Human Muc
5	36	60.0	1737	3	AAV48325 Human Muc
6	34.4	57.3	525	3	AAV48325 Human Muc
7	34.4	57.3	891	3	AAV48325 Human Muc
8	33.4	55.7	60	2	AAV48317 Nucleotid
9	31.8	53.0	60	2	AAV48318 Nucleotid
10	31.4	52.3	93	8	ABV77125 Oligonuc
11	31.4	52.3	156	8	ABV77128 Nucleotid
12	31.4	52.3	157	8	ABV77129 Nucleotid
13	31.4	52.3	157	8	ABV77127 Nucleotid
14	31.4	52.3	3621	8	ABV77131 Nucleotid
15	31.4	52.3	6877	8	ABV77133 Nucleotid
16	30.2	50.3	156	10	ADK68635 HSP65-MUC
17	30	50.0	60	2	AAV48319 Nucleotid
18	29	48.3	60	2	AAV48320 Nucleotid
19	29	48.3	60	2	AAV48324 Nucleotid
20	28.6	47.7	1800	10	ADK68629 HSP65-MUC
21	27.4	45.7	60	2	AAV48316 Nucleotid

C	22	27.4	45.7	309	1	AAV48322
C	23	27.4	45.7	1394	12	AD157712
C	24	27.4	45.7	1378	12	AD157693
C	25	27.4	45.7	1424	12	AD023180
C	26	27.4	45.7	1428	6	AB160159
C	27	27.4	45.7	1428	12	AD023125
C	28	27.4	45.7	1457	12	AD157627
C	29	27.4	45.7	1527	2	AAV48329
C	30	27.4	45.7	1614	12	ADK70370
C	31	27.4	45.7	1630	12	AD157708
C	32	27.4	45.7	1634	12	AD157689
C	33	27.4	45.7	1712	12	AD157686
C	34	27.4	45.7	1738	12	AD157669
C	35	27.4	45.7	1755	12	AD157673
C	36	27.4	45.7	1774	12	AD157691
C	37	27.4	45.7	1774	12	AD157625
C	38	27.4	45.7	1799	12	AD023124
C	39	27.4	45.7	1803	12	AD157699
C	40	27.4	45.7	1804	6	AB167539
C	41	27.4	45.7	1804	9	AAV48329
C	42	27.4	45.7	1804	10	AD14719
C	43	27.4	45.7	1804	12	AD157694
C	44	27.4	45.7	1805	12	AD028642
C	45	27.4	45.7	1808	12	AD157706

ALIGNMENTS

RESULT 1	AAV48322	strand: DNA; 60 BP.
ID	AAV48322	
XX	AAV48322	
AC	AAV48322	
XX	AAV48322	
DT	20-NOV-1998	(first entry)
XX	20-NOV-1998	
DE	Nucleotide sequence encoding MUC1 tandem repeat unit R7.	
KW	ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;	
KW	tumour; tumour-associated antigen.	
OS	Homo sapiens.	
XX	27-AUG-1998.	
XX	27-AUG-1998.	
XX	24-FEB-1998; 98WO-US003693.	
XX	24-FEB-1998; 98WO-US003693.	
XX	24-FEB-1997; 97US-0038253P.	
XX	24-FEB-1997; 97US-0038253P.	
PA	(THER-) THERION BIOLOGICS CORP.	
PA	(USSH) US DEPT HEALTH & HUMAN SERVICES.	
PA	(DAND) DNA FABER CANCER INST INC.	
PI	Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;	
DR	WPI; 1998-467492/40.	
XX	New recombinant pox virus for tumour therapy - comprises DNA encoding an	
PT	immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.	
PT	immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.	
XX	Disclosure; Page 11; 42pp; English.	
XX	The MUC1 tandem repeat units AAV48317-V48325 were used to create an	
XX	immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus	
XX	(RPV). The RPV was used in a pharmaceutical composition also containing	
XX	an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The	
XX	recombinant pox virus therefore encodes an immunogenic MUC1 fragment that	
XX	does not undergo significant genetic deletion, thereby providing an	
XX	unprecedentedly stable and immunogenic pox virus. They can be used to	
XX	prevent or treat tumours expressing MUC1 tumour-associated antigens	

XX Sequence 60 BP; 14 A; 24 C; 16 G; 6 T; 0 U; 0 Other;
SQ

Query Match 100.0%; Score 60; DB 2; Length 60;
Best Local Similarity 100.0%; Pred. No. 3.3e-11;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAAGTACCGCTCCACCTGCACACGCGGTCACACGAGCGCGCCAGACATCTGACCTGCGCCA 60
DB 1 GGAAGTACCGCTCCACCTGCACACGCGGTCACACGAGCGCGCCAGACATCTGACCTGCGCCA 60

RESULT 2
AAV48321
ID AAV48321 standard; DNA; 60 BP.

AC AAV48321;
XX
DT 20-NOV-1998 (first entry)

DE Nucleotide sequence encoding MUC1 tandem repeat unit R6.
XX
KW ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
XX tumour; tumour-associated antigen.

OS Homo sapiens.
XX
PN WO9837095-A2.

PD 27-AUG-1998.

PF 24-FEB-1998; 98WO-US003693.

PR 24-FEB-1997; 97US-0038253P.

PA (THER-) THERION BIOLOGICS CORP.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX (DAND) DANA FARBER CANCER INST INC.

PI Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;
XX
DR WPI; 1998-467492/40.

XX New recombinant pox virus for tumour therapy - comprises DNA encoding an
PT immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
XX
PS Disclosure; Page 11; 42pp; English.

XX The MUC1 tandem repeat units AAV48317-V48325 were used to create an
CC immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
CC (RPV). The RPV was used in a pharmaceutical composition also containing
CC an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes.
CC The recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
CC does not undergo significant genetic deletion, thereby providing an
CC unexpectedly stable and immunogenic pox virus. They can be used to
CC prevent or treat tumours expressing MUC1 tumour-associated antigens
XX
SQ Sequence 60 BP; 13 A; 27 C; 17 G; 3 T; 0 U; 0 Other;

Query Match 86.7%; Score 52; DB 2; Length 60;
Best Local Similarity 91.7%; Pred. No. 1.9e-08;
Matches 55; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GGAAGTACCGCTCCACCTGCACACGCGGTCACACGAGCGCGCCAGACATCTGACCTGCGCCA 60
DB 1 GGCACACCGCGCACCGCCCGGTCACACGCGGTCACACGAGCGCGCCAGACATCTGACCTGCGCCA 60

RESULT 3
AAV48325
ID AAV48325 standard; DNA; 60 BP.
XX
AC AAV48325;

XX 20-NOV-1998 (first entry)
DT

DE Nucleotide sequence encoding MUC1 tandem repeat unit R10.

KW ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
XX tumour; tumour-associated antigen.

OS Homo sapiens.

PN WO9837095-A2.

PD 27-AUG-1998.

PF 24-FEB-1998; 98WO-US003693.

PR 24-FEB-1997; 97US-0038253P.

PA (THER-) THERION BIOLOGICS CORP.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX (DAND) DANA FARBER CANCER INST INC.

PI Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;
XX
DR WPI; 1998-467492/40.

XX New recombinant pox virus for tumour therapy - comprises DNA encoding an
PT immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
XX
PS Disclosure; Page 11; 42pp; English.

XX The MUC1 tandem repeat units AAV48317-V48325 were used to create an
CC immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
CC (RPV). The RPV was used in a pharmaceutical composition also containing
CC an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes.
CC The recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
CC does not undergo significant genetic deletion, thereby providing an
CC unexpectedly stable and immunogenic pox virus. They can be used to
CC prevent or treat tumours expressing MUC1 tumour-associated antigens
XX
SQ Sequence 60 BP; 12 A; 22 C; 19 G; 7 T; 0 U; 0 Other;

Query Match 63.7%; Score 38.2; DB 2; Length 60;
Best Local Similarity 78.0%; Pred. No. 0.0011;
Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 GGAAGTACCGCTCCACCTGCACACGCGGTCACACGAGCGCGCCAGACATCTGACCTGCGGCC 59
DB 1 GGAAGTACCGCTCCACCTGCACACGCGGTCACACGAGCGCGCCAGACATCTGACCTGCGGCC 59

RESULT 4

ID AAD00388 standard; DNA; 1371 BP.

AC AAD00388;

DT 29-AUG-2000 (first entry)

DE Human Mucin 1 (MUC-1) protein fragment encoding DNA #5.

KW Human; Mucin 1; MUC-1; tumour; pMR330 expression vector; anti-tumour;
XX therapy; immune response; cytostatic; vaccine; de.

OS Homo sapiens.

PN Key Location/Qualifiers

FT CDS 1..1371

FT /*tag= a

PN /product= "MUC-1 protein fragment"

XX WO200025827-A2.

```

PD 11-MAY-2000.
XX 18-OCT-1999; 99WO-EP007874.
XX 30-OCT-1998; 98IT-MI002330.
XX (MENA ) MENARINI RICERCHE SPA.
XX Parente D, Di Massimo AM, De Santis R;
XX WPI; 2000-365410/31.
XX P-PSDB; AAY71024.
XX Composition containing one or more DNA molecules encoding fragments of a
PT Mucin 1 (MUC-1) protein overexpressed in tumor cells, useful in anti-
PT tumor therapy.
XX Claim 16; Fig 5; 56pp; English.
XX The present sequence is a DNA encoding a fragment of human Mucin 1 (MUC-
CC 1) antigenic protein which is overexpressed in tumor cells. The sequence
CC was obtained by PCR from plasmids pMRS166, pMRS167, pMRS168 and pMRS169
CC which contain MUC-1 DNA from BT20 tumor cells. It corresponds to
CC nucleotides 136-1497 of the EMBL sequence J05581 with a start codon and
CC two stop codons. The present sequence is cloned into a pMRS30 expression
CC vector and used in pharmaceutical composition e.g. vaccine for inducing
CC an antigen-specific anti-tumour immune response. Composition containing
CC this DNA molecule is useful in anti-tumour therapy of patients affected
CC with tumours characterised by high MUC-1 expression
SQ Sequence 1371 BP; 300 A; 450 C; 321 G; 300 T; 0 U; 0 Other;
Query Match 60.0%; Score 36; DB 3; Length 1371;
Best Local Similarity 75.0%; Pred. No. 0.0092;
Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 1 GGAAGTACCGCTCCACCTGACACGCGGTCACAGAAGCGCGCCAGACACTGACCTGGCGCA 60
Db 335 GGAAGTACCGCTCCACACGACACACGCGTGTACTCTCGGCTCCGAGATCCAGCGCGCCCA 384
RESULT 5
AAD00394
ID AAD00394 standard; DNA; 1737 BP.
XX
XX AAD00394;
AC 15-SEP-2003 (revised)
DT 29-AUG-2000 (first entry)
XX
XX Ubiquitin-E. coli LacI-human Mucin 1 fusion protein encoding DNA #5.
XX
XX Ubiquitin; LacI; beta-galactosidase; fusion protein; human; Mucin 1;
KM MUC-1; tumour; pMRS30 expression vector; anti-tumour; therapy;
KW immune response; cytostatic; vaccine; ds.
XX Homo sapiens.
OS Escherichia coli.
OS Chimeric.
XX
XX Key Location/Qualifiers
XX CDS 1..1737
XX FT /*tag= a
XX FT /product= "Ubiquitin-LacI-MUC-1 fusion protein"
XX FT 1..369
XX FT /*tag= b
XX FT /label= UBILacI DNA
XX FT /note= "includes ubiquitin-E. coli LacI fusion DNA"
XX FT 370..1737
XX FT /*tag= c
XX FT /note= "Human MUC-1 partial DNA that corresponds to
XX FT nucleotides 136-1497 of the EMBL sequence J05581 with two
XX FT stop codons"

```

```

XX WO200025827-A2.
XX 11-MAY-2000.
XX 18-OCT-1999; 99WO-EP007874.
XX 30-OCT-1998; 98IT-MI002330.
XX (MENA ) MENARINI RICERCHE SPA.
XX Parente D, Di Massimo AM, De Santis R;
XX WPI; 2000-365410/31.
XX P-PSDB; AAY71030.
XX Composition containing one or more DNA molecules encoding fragments of a
PT Mucin 1 (MUC-1) protein overexpressed in tumor cells, useful in anti-
PT tumor therapy.
XX Claim 18; Fig 11; 56pp; English.
XX The present sequence is a DNA encoding a fusion protein consisting of
CC human Mucin 1 (MUC-1) fragment fused to UBILacI sequence at the N-
CC terminus. The UBILacI sequence consists of ubiquitin from MCF7 cell line
CC and a portion of E. coli beta-galactosidase (LacI). MUC-1 is an antigenic
CC protein overexpressed in tumor cells. The present sequence is cloned
CC into a pMRS30 expression vector and used in pharmaceutical composition
CC e.g. vaccine for inducing an antigen-specific anti-tumour immune
CC response. Composition containing this DNA molecule is useful in anti-
CC tumour therapy of patients affected with tumours characterised by high
CC MUC-1 expression. (Updated on 15-SEP-2003 to standardise OS field)
XX
SQ Sequence 1737 BP; 392 A; 558 C; 418 G; 369 T; 0 U; 0 Other;
Query Match 60.0%; Score 36; DB 3; Length 1737;
Best Local Similarity 75.0%; Pred. No. 0.0094;
Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 1 GGAAGTACCGCTCCACCTGACACGCGGTCACAGAAGCGCGCCAGACACTGACCTGGCGCA 60
Db 691 GGAAGTACCGCTCCACACGACACACGCGTGTACTCTCGGCTCCGAGATCCAGCGCGCCCA 750
RESULT 6
AAD00385
ID AAD00385 standard; DNA; 525 BP.
XX
XX AAD00385;
AC 29-AUG-2000 (first entry)
DT 29-AUG-2000 (first entry)
XX
XX Human Mucin 1 (MUC-1) protein fragment encoding DNA #2.
XX
XX Human; Mucin 1; tumour; pMRS30 expression vector; anti-tumour;
KM therapy; immune response; cytostatic; vaccine; ds.
XX Homo sapiens.
OS
OS Key Location/Qualifiers
OS CDS 1..525
OS FT /*tag= a
OS FT /product= "MUC-1 protein fragment"
XX
XX WO200025827-A2.
XX 11-MAY-2000.
XX 18-OCT-1999; 99WO-EP007874.
XX 30-OCT-1998; 98IT-MI002330.
XX (MENA ) MENARINI RICERCHE SPA.

```

```
XX
PI Parente D, Di Massimo AM, De Santis R;
XX WPI; 2000-365410/31.
DR P-PSDB; AAY71021.
XX
PT Composition containing one or more DNA molecules encoding fragments of a
PT Mucin 1 (MUC-1) protein overexpressed in tumor cells, useful in anti-
PT tumor therapy.
XX
PS Claim 16; Fig 2; 56pp; English.
XX
CC The present sequence is a DNA encoding a fragment of human Mucin 1 (MUC-
CC 1) antigenic protein which is overexpressed in tumor cells. The sequence
CC was obtained from BT20 tumour cells by reverse transcriptase-PCR and
CC corresponds to nucleotides 205-720 of the EMBL sequence J05581 with a
CC start codon and two stop codons. The present sequence is cloned into a
CC pMR330 expression vector and used in pharmaceutical composition e.g.
CC vaccine for inducing an antigen-specific anti-tumour immune response.
CC Composition containing this DNA molecule is useful in anti-tumour therapy
CC of patients affected with tumours characterised by high MUC-1 expression
XX
SQ Sequence 525 BP; 111 A; 199 C; 128 G; 87 T; 0 U; 0 Other;
Query Match
Best Local Similarity 57.3%; Score 34.4; DB 3; Length 525;
Matches 44; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
XX
QY 1 GGAAGTACCGCTCCACCTGCACACGGGGGTCAAGCGCGGACACTCGACCTGCGCCA 60
DB 256 GGAAGTACCTGCTCCACCGACACGAGTGTACTCTGCGCTCGGATACACAGCGCGGCCCA 315
RESULT 7
AAB00391
ID AAB00391 standard; DNA; 891 BP.
XX
AC AAB00391;
XX
DT 15-SEP-2003 (revised)
DT 29-AUG-2000 (first entry)
XX
DE Ubiquitin-E. coli LacI-human Mucin 1 fusion protein encoding DNA #2.
XX
KW Ubiquitin; LacI; beta-galactosidase; fusion protein; human; Mucin 1;
KW MUC-1; tumour; pMR330 expression vector; anti-tumour; therapy;
XX
OS Homo sapiens.
OS Escherichia coli.
OS Chimeric.
XX
FH Key
FH CDS
FT location/Qualifiers
FT 1..891
FT /*tag= a
FT /product= "Ubiquitin-LacI-MUC-1 fusion protein"
FT 1..369
FT /*tag= b
FT /label= UBILacI DNA
FT /notes= "Includes ubiquitin-E. coli LacI fusion DNA"
FT 370..891
FT /*tag= c
FT /note= "Human MUC-1 partial DNA that corresponds to
FT nucleotides 205-720 of the EMBL sequence J05581 with two
FT stop codons"
XX
FN WO200025827-A2.
XX
PD 11-MAY-2000.
XX
PF 18-OCT-1999; 99WO-EP007874.
XX
PR 30-OCT-1998; 98IT-MI002330.
```

```
XX
PA (MENA ) MENARINI RICERCHE SPA.
XX
PI Parente D, Di Massimo AM, De Santis R;
XX WPI; 2000-365410/31.
DR P-PSDB; AAY71027.
XX
PT Composition containing one or more DNA molecules encoding fragments of a
PT Mucin 1 (MUC-1) protein overexpressed in tumor cells, useful in anti-
PT tumor therapy.
XX
PS Claim 16; Fig 8; 56pp; English.
XX
CC The present sequence is a DNA encoding a fusion protein consisting of
CC human Mucin 1 (MUC-1) fragment fused to Ubilacti sequence at the N-
CC terminus. The Ubilacti sequence consists of ubiquitin from MCF7 cell line
CC and a portion of E. coli beta-galactosidase (lacI). MUC-1 is an antigenic
CC protein overexpressed in tumour cells. The present sequence is cloned
CC into a pMR330 expression vector and used in pharmaceutical composition
CC e.g. vaccine for inducing an antigen-specific anti-tumour immune
CC response. Composition containing this DNA molecule is useful in anti-
CC tumour therapy of patients affected with tumours characterised by high
CC MUC-1 expression. (Updated on 15-SEP-2003 to standardise OS field)
XX
SQ Sequence 891 BP; 203 A; 307 C; 225 G; 156 T; 0 U; 0 Other;
Query Match
Best Local Similarity 57.3%; Score 34.4; DB 3; Length 891;
Matches 44; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
XX
QY 1 GGAAGTACCGCTCCACCTGCACACGGGGGTCAAGCGCGGACACTCGACCTGCGCCA 60
DB 622 GGAAGTACCTGCTCCACCGACACGAGTGTACTCTGCGCTCGGATACACAGCGCGGCCCA 681
RESULT 8
AAV48317
ID AAV48317 standard; DNA; 60 BP.
XX
AC AAV48317;
XX
DT 20-NOV-1998 (first entry)
XX
DE Nucleotide sequence encoding MUC1 tandem repeat unit R2.
XX
KW 86; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
KW tumour; tumour-associated antigen.
XX
OS Homo sapiens.
XX
FN WO9837095-A2.
XX
PD 27-AUG-1998.
XX
PF 24-FEB-1998; 98WO-US003693.
XX
PR 24-FEB-1997; 97US-0038253P.
XX
PA (THER-) THERION BIOLOGICS CORP.
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA (DAND ) DANA FARBER CANCER INST INC.
XX
PI Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;
XX WPI; 1998-467492/40.
XX
PT New recombinant pox virus for tumour therapy - comprises DNA encoding an
PT immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
XX
PS Disclosure; Page 11; 42pp; English.
XX
CC The MUC1 tandem repeat units AAV48317-V48325 were used to create an
```


CC immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus (RPV). The RPV was used in a pharmaceutical composition also containing CC an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The CC recombinant pox virus therefore encodes an immunogenic MUC1 fragment that CC does not undergo significant genetic deletion, thereby providing an CC unexpectedly stable and immunogenic pox virus. They can be used to CC prevent or treat tumours expressing MUC1 tumour-associated antigens XX

Sequence 60 BP; 17 A; 21 C; 13 G; 9 T; 0 U; 0 Other;

Query Match 55.7%; Score 33.4; DB 2; Length 60;
Best Local Similarity 72.9%; Pred. No. 0.05;
Matches 43; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1 GGAAGTACCGCTCCACCTGACACGCGGCTCACAAGCGCCGACACTGCTGCGCC 59
DB 1 GGCAGTACTGCACACCGGACATGGCGTAACTAGACACTGATACAGACTCTGCACC 59

RESULT 9

AAV48318
ID AAV48318 standard; DNA; 60 BP.

AC AAV48318;

DT 20-NOV-1998 (first entry)

DE Nucleotide sequence encoding MUC1 tandem repeat unit R3.

KM 88; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
tumour; tumour-associated antigen.

OS Homo sapiens.

PN WO9837095-A2.

PD 27-AUG-1998.

PF 24-FEB-1998; 98WO-US003693.

PR 24-FEB-1997; 97US-0038253P.

PA (THER-) THERION BIOLOGICS CORP.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

PA (DAND) DANA FARBER CANCER INST INC.

PI Schlom J, Kantor J, Kufe D, Panicali D, Grtitz L;
MPI; 1998-467492/40.

PT New recombinant pox virus for tumour therapy - comprises DNA encoding an
immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.

PS Disclosure; Page 11; 42pp; English.

CC The MUC1 tandem repeat units AAV48317-V48325 were used to create an
immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
(RPV). The RPV was used in a pharmaceutical composition also containing
an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The
CC recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
CC does not undergo significant genetic deletion, thereby providing an
unexpectedly stable and immunogenic pox virus. They can be used to
prevent or treat tumours expressing MUC1 tumour-associated antigens XX

Sequence 60 BP; 7 A; 29 C; 19 G; 5 T; 0 U; 0 Other;

Query Match 53.0%; Score 31.8; DB 2; Length 60;
Best Local Similarity 71.2%; Pred. No. 0.18;
Matches 42; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 1 GGAAGTACCGCTCCACCTGACACGCGGCTCACAAGCGCCGACACTGCTGCGCC 59
DB 1 GGATCACCAGCGCGCTCTGCGACGAGTACGCTCGGCCGCCGACACGCGCCGCTCC 59

RESULT 10

ABV77125
ID ABV77125 standard; DNA; 93 BP.

AC ABV77125;

DT 28-MAR-2003 (first entry)

DE Oligonucleotide used to create a plasmid insert encoding MUC1 epitope.

KM Intracellular analysis; intracellular interaction; intrabody; ScFv; MUC1;
mucin 1; ss.

OS Synthetic.

PN WO200286505-A2.

PD 31-OCT-2002.

PF 02-APR-2002; 2002WO-GB001235.

PR 31-MAR-2001; 2001GB-00008165.

PA (UYMA-) UNIV VICTORIA MANCHESTER.

PI Roderick SPB;

DR MPI; 2003-103417/09.

PT Intracellularly analyzing or detecting the presence of a target molecule
within a biological cell, by determining the presence or signal
representative of binding of polypeptide species to the target molecule.

PS Example; Page 19; 51pp; English.

CC The specification describes a method for the intracellular analysis of a
CC target molecule within a biological cell. The method comprises expressing
CC within the cell a first polypeptide sequence comprised of a first binding
CC species capable of binding to the target molecule and a first reporter
CC moiety attached to the first binding species; expressing within the cell
CC a second polypeptide sequence comprised of a second binding species
CC capable of competing with the target molecule for binding of the first binding
CC of competing with the target molecule for binding of the first binding
CC species and a second reporter moiety, the first and second reporter
CC moieties being such that on binding together of the first and second
CC binding species the first and second reporter moieties interact so as to
CC be capable of producing a signal that can be differentiated from one
CC capable of being generated when the first and second reporter moieties do
CC not interact; and effecting a measurement to determine the presence or
CC signal representative of binding of the first and second binding species.
CC The method is useful for the intracellular analysis of a target molecule,
CC e.g. for detecting the presence and/or amount of target molecules in
CC cells. The method is also useful for studying the interactions of
CC intracellular molecules. Oligonucleotides ABV77125-26 were used to create
CC a plasmid insert encoding the MUC1 epitope of human mucin 1. The
CC oligonucleotides were used in an intracellular assay for MUC1,
CC demonstrating the method of the invention. In this assay, the first
CC polypeptide is an anti-MUC1 ScFv, and the second polypeptide comprises
CC the MUC1 epitope attached to yellow fluorescent protein XX

Sequence 93 BP; 20 A; 41 C; 19 G; 13 T; 0 U; 0 Other;

Query Match 52.3%; Score 31.4; DB 8; Length 93;
Best Local Similarity 85.4%; Pred. No. 0.26;
Matches 35; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GGAAGTACCGCTCCACCTGACACGCGGCTCACAAGCGCGCC 41
DB 38 GGATCTACCGCTCTCTGCGACGAGTACGAGTACAGCGCACCC 78

```
RESULT 11
ABV77128
ID ABV77128 standard; DNA; 156 BP.
XX
AC ABV77128;
XX
DT 28-MAR-2003 (first entry)
XX
DE Nucleotide sequence of a mutated plasmid insert encoding MUC1 epitope.
XX
KM Intracellular analysis; intracellular interaction; intrabody; scFv; MUC1;
KM mucin 1; ss.
XX
OS Synthetic.
XX
FT Key Location/Qualifiers
FT 11.115
FT CDS /*tag= a
FT /note= "partial sequence"
XX
PN MO200286505-A2.
XX
PD 31-OCT-2002.
XX
PE 02-APR-2002; 2002MO-GB001235.
XX
PR 31-MAR-2001; 2001GB-00008165.
XX
PA (UYMA-) UNIV VICTORIA MANCHESTER.
XX
PI Roderick SPB;
XX
DR MPI; 2003-103417/09.
DR P-PSDB; ABB99633.
XX
PT Intracellularly analyzing or detecting the presence of a target molecule
PT within a biological cell, by determining the presence or signal
PT representative of binding of polypeptide species to the target molecule.
XX
XX
PS Example; Page 20-21; 51pp; English.
XX
CC The specification describes a method for the intracellular analysis of a
CC target molecule within a biological cell. The method comprises expressing
CC within the cell a first polypeptide sequence comprised of a first binding
CC species capable of binding to the target molecule and a first reporter
CC moiety attached to the first binding species; expressing within the cell
CC a second polypeptide sequence comprised of a second binding species
CC capable of competing with the target molecule for binding within the cell
CC of competing with the target molecule for binding of the first binding
CC species and a second reporter moiety, the first and second reporter
CC moieties being such that on binding together of the first and second
CC binding species the first and second reporter moieties interact so as to
CC be capable of producing a signal that can be differentiated from one
CC capable of being generated when the first and second reporter moieties do
CC not interact; and effecting a measurement to determine the presence or
CC signal representative of binding of the first and second binding species.
CC The method is useful for the intracellular analysis of a target molecule,
CC e.g. for detecting the presence and/or amount of target molecules in
CC cells. The method is also useful for studying the interactions of
CC intracellular molecules. The present sequence encodes a mutated MUC1
CC epitope of human mucin 1 and its surrounding amino acids in a tandem
CC repeat. It was used in an intracellular assay for MUC1, demonstrating the
CC method of the invention. In this assay, the first polypeptide is an anti-
CC MUC1 scFv, and the second polypeptide comprises the MUC1 epitope attached
CC to yellow fluorescent protein
XX
SQ Sequence 156 BP; 32 A; 69 C; 33 G; 22 T; 0 U; 0 Other;
```

```
Query Match 52.3%; Score 31.4; DB 8; Length 156;
Best Local Similarity 85.4%; Pred. No. 0.27;
Matches 35; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
1 GGAAGTACCGCTCCACCTGCACACGCGGTCAACAACGCGCC 41
```

```
Db
38 GGATCTACCGCTCTCTGTCACCGAGTCAACAACGCGACC 78
|||||
RESULT 12
ABV77129
ID ABV77129 standard; DNA; 157 BP.
XX
AC ABV77129;
XX
DT 28-MAR-2003 (first entry)
XX
DE Nucleotide sequence of a plasmid insert encoding MUC1 epitope.
XX
KM Intracellular analysis; intracellular interaction; intrabody; scFv; MUC1;
KM mucin 1; ss.
XX
OS Synthetic.
XX
FT Key Location/Qualifiers
FT 11.157
FT CDS /*tag= a
FT /note= "partial sequence"
XX
PN MO200286505-A2.
XX
PD 31-OCT-2002.
XX
PE 02-APR-2002; 2002MO-GB001235.
XX
PR 31-MAR-2001; 2001GB-00008165.
XX
PA (UYMA-) UNIV VICTORIA MANCHESTER.
XX
PI Roderick SPB;
XX
DR MPI; 2003-103417/09.
DR P-PSDB; ABB99634.
XX
PT Intracellularly analyzing or detecting the presence of a target molecule
PT within a biological cell, by determining the presence or signal
PT representative of binding of polypeptide species to the target molecule.
XX
XX
PS Example; Page 21-22; 51pp; English.
XX
CC The specification describes a method for the intracellular analysis of a
CC target molecule within a biological cell. The method comprises expressing
CC within the cell a first polypeptide sequence comprised of a first binding
CC species capable of binding to the target molecule and a first reporter
CC moiety attached to the first binding species; expressing within the cell
CC a second polypeptide sequence comprised of a second binding species
CC capable of competing with the target molecule for binding within the cell
CC of competing with the target molecule for binding of the first binding
CC species and a second reporter moiety, the first and second reporter
CC moieties being such that on binding together of the first and second
CC binding species the first and second reporter moieties interact so as to
CC be capable of producing a signal that can be differentiated from one
CC capable of being generated when the first and second reporter moieties do
CC not interact; and effecting a measurement to determine the presence or
CC signal representative of binding of the first and second binding species.
CC The method is useful for the intracellular analysis of a target molecule,
CC e.g. for detecting the presence and/or amount of target molecules in
CC cells. The method is also useful for studying the interactions of
CC intracellular molecules. The present sequence encodes the MUC1 epitope of
CC human mucin 1 and its surrounding amino acids in a tandem repeat. It was
CC used in an intracellular assay for MUC1, demonstrating the method of the
CC invention. In this assay, the first polypeptide is an anti-MUC1 scFv, and
CC the second polypeptide comprises the MUC1 epitope attached to yellow
CC fluorescent protein
XX
SQ Sequence 157 BP; 32 A; 68 C; 35 G; 22 T; 0 U; 0 Other;
```

```
Query Match 52.3%; Score 31.4; DB 8; Length 157;
```


Query Match Best Local Similarity 52.3%; Score 31.4; DB 8; Length 3621;
 Matches 35; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 1 GGAAGTACCGCTCCACCTGCACACGCGGGTCAACAAGCGGCC 41
 DB 659 GGATCTACCGCTCCTCCTCCGACGAGTCAACAAGCGCAC 699

RESULT 15

ABV77133
 ID ABV77133 standard; DNA; 6877 BP.

AC ABV77133;

DT 28-MAR-2003 (first entry)

DE Nucleotide sequence of plasmid pBudMUC-BYFPscFVECFP.

KM Intracellular analysis; intracellular interaction; intrabody; ScFv; MUC1,
 mucin 1; pBudMUC-BYFPscFVECFP; ss.

OS Synthetic.

PN WO200286505-A2.

PD 31-OCT-2002.

PF 02-APR-2002; 2002WO-GB001235.

PR 31-MAR-2001; 2001GB-00008165.

PA (UWMA-) UNIV VICTORIA MANCHESTER.

PI Roderick SPB;

DR WPI; 2003-103417/09.

PT Intracellularly analyzing or detecting the presence of a target molecule
 within a biological cell, by determining the presence or signal
 representative of binding of polypeptide species to the target molecule.

PS Example; Page 37-40; 51pp; English.

XX The specification describes a method for the intracellular analysis of a
 CC target molecule within a biological cell. The method comprises expressing
 CC within the cell a first polypeptide sequence comprised of a first binding
 CC species capable of binding to the target molecule and a first reporter
 CC moiety attached to the first binding species; expressing within the cell
 CC a second polypeptide sequence comprised of a second binding species
 CC capable of competing with the target molecule for binding species
 CC of competing with the target molecule for binding of the first binding
 CC species and a second reporter moiety, the first and second reporter
 CC moieties being such that on binding together of the first and second
 CC binding species the first and second reporter moieties interact so as to
 CC be capable of producing a signal that can be differentiated from one
 CC not interact; and effecting a measurement to determine the presence or
 CC signal representative of binding of the first and second binding species.
 CC The method is useful for the intracellular analysis of a target molecule,
 CC e.g. for detecting the presence and/or amount of target molecules in
 CC intracellular molecules. The present sequence represents a plasmid used
 CC in an intracellular assay for the MUC1 epitope of human mucin1,
 CC demonstrating the method of the invention. In this assay, the first
 CC polypeptide is an anti-MUC1 ScFv, and the second polypeptide comprises
 CC the MUC1 epitope attached to yellow fluorescent protein
 XX

Sequence 6877 BP; 1511 A; 1962 C; 1964 G; 1440 T; 0 U; 0 Other;

Query Match

Best Local Similarity 52.3%; Score 31.4; DB 8; Length 6877;

Matches 35; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GGAAGTACCGCTCCACCTGCACACGCGGGTCAACAAGCGGCC 41
 DB 701 GGATCTACCGCTCCTCCTCCGACGAGTCAACAAGCGCAC 741

Search completed: January 15, 2005, 20:36:11
 Job time : 171.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using SW model

Run on: January 15, 2005, 17:40:52 : Search time 722.8 Seconds
(without alignment)
3925.542 Million cell updates/sec

Title: US-10-057-136-9

Perfect score: 60
Sequence: 1 GGAAGTACGCTCCACCTGC.....CAGACACTCGACTGCGCCA 60

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:
1: gb_bai:*
2: gb_hgt:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pac:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_str:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	2297	6	BD272907 A recombinant
2	36	60.0	1371	6	BD225144 Medicinal
3	36	60.0	1737	6	BD225150 Medicinal
4	34.4	57.3	525	6	BD225147 Medicinal
5	34.4	57.3	891	6	BD225147 Medicinal
6	32.8	54.7	132	6	AX180276 Sequence
7	32.8	54.7	144	6	AX180276 Sequence
8	31.4	52.3	93	6	AX703425 Sequence
9	31.4	52.3	156	6	AX703427 Sequence
10	31.4	52.3	157	6	AX703428 Sequence
11	31.4	52.3	157	6	AX703431 Sequence
12	31.4	52.3	3621	6	AX703433 Sequence
13	31.4	52.3	6877	6	AX703435 Sequence
14	31.2	52.0	132	6	AX180282 Sequence
15	31.2	52.0	144	6	AX180281 Sequence
16	30	50.0	736	10	AF484219 Cavia por
17	29	48.3	303400	1	AP006578 Gloebact
18	28	46.7	32748	1	AB070951 Streptomy
19	28	46.7	296300	1	AP005035 Streptomy

C	20	27.4	45.7	120	6	AX192396	AX192396 Sequence
C	21	27.4	45.7	120	6	BD000571	BD000571 Human pol
C	22	27.4	45.7	1414	12	AF423031	AF423031 Synthetic
C	23	27.4	45.7	1455	6	CQ715242	CQ715242 Sequence
C	24	27.4	45.7	1457	6	AX959914	AX959914 Sequence
C	25	27.4	45.7	1774	6	AX959684	AX959684 Sequence
C	26	27.4	45.7	1774	6	AX959912	AX959912 Sequence
C	27	27.4	45.7	1804	6	AR492318	AR492318 Sequence
C	28	27.4	45.7	1804	6	AX335367	AX335367 Sequence
C	29	27.4	45.7	1804	6	HDMUCAB	J05581 Human polyom
C	30	27.4	45.7	1834	12	AF423030	AF423030 Synthetic
C	31	27.4	45.7	1835	6	AX959918	AX959918 Sequence
C	32	27.4	45.7	2135	6	AX959916	AX959916 Sequence
C	33	27.4	45.7	4139	6	C0834017	C0834017 Sequence
C	34	27.4	45.7	4139	6	AX334889	AX334889 Sequence
C	35	27.4	45.7	4139	6	AX335372	AX335372 Sequence
C	36	27.4	45.7	4139	6	AX336712	AX336712 Sequence
C	37	27.4	45.7	4139	6	AX409474	AX409474 Sequence
C	38	27.4	45.7	4139	6	AX440481	AX440481 Sequence
C	39	27.4	45.7	4139	6	HDMPANMU	J05582 Human pancr
C	40	27.4	45.7	7188	9	AY463543	AY463543 Homo sapi
C	41	27.4	45.7	8181	6	AX406624	AX406624 Sequence
C	42	27.4	45.7	8181	9	HDMPEM	M61170 Human polyom
C	43	27.4	45.7	8186	6	AR492320	AR492320 Sequence
C	44	27.4	45.7	133525	9	AL713399	AL713399 Human DNA
C	45	26.8	44.7	216050	1	AL646076	AL646076 Ralstonia

ALIGNMENTS

RESULT 1	BD272907	2297 bp	DNA	linear	PAT 17-JUL-2003
LOCUS	BD272907				
DEFINITION	A recombinant vector expressing multiple constitutively molecules and uses thereof.				
ACCESSION	BD272907.1	GI:33082675			
VERSION	JP 2002531133-A/1.				
KEYWORDS	synthetic construct				
SOURCE	artificial sequences.				
ORGANISM	1 (bases 1 to 2297)				
REFERENCE	Schlom,J., Hodge,J.F. and Panicali,D.				
AUTHORS	and uses thereof				
TITLE	Patent: JP 2002531133-A 1 24-SEP-2002;				
JOURNAL	THE UNITED STATES OF AMERICA, THERION BIOLOGICS CORP				
OS	Artificial Sequence				
COMMENT	PN JP 2002531133-A/1				
	PD 24-SEP-2002				
	PF 12-NOV-1999 JP 2000586927				
	PR 09-DEC-1998 US 60/111582				
	PI JEFFREY SCHLOM, JAMES HODGE, DENNIS PANICALI				
	PC C12N15/02, A61K39/12, A61K35/12, A61K35/74, A61K35/76, A61K38/00, A61K39/00,				
	PC A61K39/12, A61K39/21, A61K39/235, A61K39/245, A61K39/275, A61K39/				
	PC A61K39/125, A61K39/21, A61K39/235, A61K39/245, A61K39/275, A61K39/				
	PC 29, A61K48/00,				
	PC A61P1/04, A61P31/04, A61P31/10, A61P31/12, A61P35/00, A61P37/02, PC				
	A61P37/06,				
	PC C12N1/15, C12N1/19, C12N1/21, C12N5/10, C12N7/00, C12Q1/02, G01N3/				
	PC 53, C12N15/00,				
	PC C12N5/00, A61K37/02				
	CC Description of Artificial Sequence: VECTOR SEQUENCE FH Key				
	Location/Qualifiers				
	FT source 1..2297 /organism='Artificial Sequence'.				
FEATURES	Location/Qualifiers				
source	1..2297				
	/organism="synthetic construct"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:33630"				
ORIGIN					

Query Match 100.0%; Score 60; DB 6; Length 2297;
 Best Local Similarity 100.0%; Pred. No. 1e-09;
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 GGAAGTACCGCTCCACCTGCAACGCGGCTCACAAGCGCGCAGACACTGACCTGGCCCA 60
 706 GGAAGTACCGCTCCACCTGCAACGCGGCTCACAAGCGCGCAGACACTGACCTGGCCCA 765

RESULT 2
 LOCUS BD225144 1371 bp DNA linear PAT 17-JUL-2003
 DEFINITION Medicinal composition having antitumor effect and containing DNA
 ACCESSION BD225144
 VERSION BD225144.1 GI:33034914
 KEYWORDS JP 2002528519-A/5.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 Pallette,D., Massimo,A.M.D. and Desantis,R.
 1 (bases 1 to 1371)
 encoding antigenic protein
 Medicinal composition having antitumor effect and containing DNA
 Patent: JP 2002528519-A 5 03-SEP-2002;

COMMENT
 MENARINI RICERCHE SPA
 OS Homo sapiens (human)
 PN JP 2002528519-A/5
 PD 03-SEP-2002
 PF 18-OCT-1999 JP 2000579265
 PR 30-OCT-1998 IT MI98A002330
 PI DINO PALLEENTE,ANNA MARIA D MASSIMO,RITA DESANTIS PC
 A61K38/00,A61K35/76,A61K39/00,A61K48/00,A61P35/00,C12N15/09, PC
 A61K37/02,
 PC C12N15/00
 CC Medicinal composition having antitumor effect and containing
 CC antigenic protein
 CC DNA encoding
 FH Key Location/Qualifiers
 FT source 1..1371
 /organism='Homo sapiens (human)'.
 Location/Qualifiers
 1..1371
 /organism='Homo sapiens'
 /mol_type='genomic DNA'
 /db_xref='taxon:9606'

FEATURES
 source
 1..1371
 /organism='Homo sapiens'
 /mol_type='genomic DNA'
 /db_xref='taxon:9606'

ORIGIN
 Query Match 60.0%; Score 36; DB 6; Length 1371;
 Best Local Similarity 75.0%; Pred. No. 0.14;
 Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Db 1 GGAAGTACCGCTCCACCTGCAACGCGGCTCACAAGCGCGCAGACACTGACCTGGCCCA 60
 325 GGAAGTACCGCTCCACCTGCAACGCGGCTCACAAGCGCGCAGACACTGACCTGGCCCA 384

RESULT 3
 LOCUS BD225150 1737 bp DNA linear PAT 17-JUL-2003
 DEFINITION Medicinal composition having antitumor effect and containing DNA
 ACCESSION BD225150
 VERSION BD225150.1 GI:33034920
 KEYWORDS JP 2002528519-A/11.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 Pallette,D., Massimo,A.M.D. and Desantis,R.
 1 (bases 1 to 1737)
 encoding antigenic protein
 Medicinal composition having antitumor effect and containing DNA
 Patent: JP 2002528519-A 2 03-SEP-2002;

COMMENT
 MENARINI RICERCHE SPA
 OS Homo sapiens (human)
 PN JP 2002528519-A/2
 PD 03-SEP-2002
 PF 18-OCT-1999 JP 2000579265
 PR 30-OCT-1998 IT MI98A002330
 PI DINO PALLEENTE,ANNA MARIA D MASSIMO,RITA DESANTIS PC
 A61K38/00,A61K35/76,A61K39/00,A61K48/00,A61P35/00,C12N15/09, PC
 A61K37/02,
 PC C12N15/00
 CC Medicinal composition having antitumor effect and containing
 CC antigenic protein
 CC DNA encoding
 FH Key Location/Qualifiers
 FT source 1..1737
 /organism='Homo sapiens (human)'.
 Location/Qualifiers
 1..1737
 /organism='Homo sapiens'
 /mol_type='genomic DNA'
 /db_xref='taxon:9606'

FEATURES
 source
 1..1737
 /organism='Homo sapiens (human)'.
 Location/Qualifiers
 1..1737
 /organism='Homo sapiens'
 /mol_type='genomic DNA'
 /db_xref='taxon:9606'

TITLE
 JOURNAL
 COMMENT
 MEDICAL COMPOSITION HAVING ANTITUMOR EFFECT AND CONTAINING DNA
 encoding antigenic protein
 Patent: JP 2002528519-A 11 03-SEP-2002;
 MENARINI RICERCHE SPA
 OS Homo sapiens (human)
 PN JP 2002528519-A/11
 PD 03-SEP-2002
 PF 18-OCT-1999 JP 2000579265
 PR 30-OCT-1998 IT MI98A002330
 PI DINO PALLEENTE,ANNA MARIA D MASSIMO,RITA DESANTIS PC
 A61K38/00,A61K35/76,A61K39/00,A61K48/00,A61P35/00,C12N15/09, PC
 A61K37/02,
 PC C12N15/00
 CC Medicinal composition having antitumor effect and containing
 CC antigenic protein
 CC DNA encoding
 FH Key Location/Qualifiers
 FT source 1..1737
 /organism='Homo sapiens (human)'.
 Location/Qualifiers
 1..1737
 /organism='Homo sapiens'
 /mol_type='genomic DNA'
 /db_xref='taxon:9606'

FEATURES
 source
 1..1737
 /organism='Homo sapiens (human)'.
 Location/Qualifiers
 1..1737
 /organism='Homo sapiens'
 /mol_type='genomic DNA'
 /db_xref='taxon:9606'

ORIGIN
 Query Match 60.0%; Score 36; DB 6; Length 1737;
 Best Local Similarity 75.0%; Pred. No. 0.13;
 Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Db 1 GGAAGTACCGCTCCACCTGCAACGCGGCTCACAAGCGCGCAGACACTGACCTGGCCCA 60
 691 GGAAGTACCGCTCCACCTGCAACGCGGCTCACAAGCGCGCAGACACTGACCTGGCCCA 750

RESULT 4
 LOCUS BD225141 525 bp DNA linear PAT 17-JUL-2003
 DEFINITION Medicinal composition having antitumor effect and containing DNA
 ACCESSION BD225141
 VERSION BD225141.1 GI:33034911
 KEYWORDS JP 2002528519-A/2.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 Pallette,D., Massimo,A.M.D. and Desantis,R.
 1 (bases 1 to 525)
 encoding antigenic protein
 Medicinal composition having antitumor effect and containing DNA
 Patent: JP 2002528519-A 2 03-SEP-2002;

COMMENT
 MENARINI RICERCHE SPA
 OS Homo sapiens (human)
 PN JP 2002528519-A/2
 PD 03-SEP-2002
 PF 18-OCT-1999 JP 2000579265
 PR 30-OCT-1998 IT MI98A002330
 PI DINO PALLEENTE,ANNA MARIA D MASSIMO,RITA DESANTIS PC
 A61K38/00,A61K35/76,A61K39/00,A61K48/00,A61P35/00,C12N15/09, PC
 A61K37/02,
 PC C12N15/00
 CC Medicinal composition having antitumor effect and containing
 CC antigenic protein
 CC DNA encoding
 FH Key Location/Qualifiers
 FT source 1..525
 /organism='Homo sapiens (human)'.
 Location/Qualifiers
 1..525
 /organism='Homo sapiens'
 /mol_type='genomic DNA'
 /db_xref='taxon:9606'

FEATURES
 source
 1..525
 /organism='Homo sapiens (human)'.
 Location/Qualifiers
 1..525
 /organism='Homo sapiens'
 /mol_type='genomic DNA'
 /db_xref='taxon:9606'

ORIGIN

Query Match 57.3%; Score 34.4; DB 6; Length 525;
Best Local Similarity 73.3%; Pred. No. 0.52;
Matches 44; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 1 GGAAGTACCGCTCCACCTGCACACGCGGTTCACAAAGCGCCGACACCTGCAGCTGGGCCA 60
|||||
Db 256 GGAAGTACCTGCTCCACACGACACACGCGTGTACCTCGGCTCCGATACACAGCGCCGCCCA 315

RESULT 5
BD225147
LOCUS BD225147 891 bp DNA linear PAT 17-JUL-2003
DEFINITION Medicinal composition having antitumor effect and containing DNA

ACCESSION BD225147
BD225147
VERSION JP 2002528519-A/8.
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
TITLE 1 (bases 1 to 891)
Pallente, D., Massimo, A.M.D. and Desantis, R.
JOURNAL Medicinal composition having antitumor effect and containing DNA

COMMENT
Patent: JP 2002528519-A 8 03-SEP-2002;
MENARINI RICERCHE SPA
OS Homo sapiens (human)
PN JP 2002528519-A/8
PD 03-SEP-2002
PF 18-OCT-1999 JP 2000579265
PR 30-OCT-1998 IT M198A002330
PI DINO PALLENTE, ANNA MARIA D MASSIMO, RITA DESANTIS PC
A61K38/00, A61K35/76, A61K39/00, A61K48/00, A61P35/00, C12N15/09, PC
A61K37/02,
PC C12N15/00
CC Medicinal composition having antitumor effect and containing
CC antigenic protein
CC DNA encoding
FH Key Location/Qualifiers
FT source 1. 891
Location/Qualifiers
1. 891
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

FEATURES

source

ORIGIN

Query Match 57.3%; Score 34.4; DB 6; Length 891;
Best Local Similarity 73.3%; Pred. No. 0.49;
Matches 44; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 1 GGAAGTACCGCTCCACCTGCACACGCGGTTCACAAAGCGCCGACACCTGCAGCTGGGCCA 60
|||||
Db 622 GGAAGTACCTGCTCCACACGACACACGCGTGTACCTCGGCTCCGATACACAGCGCGGCCCA 681

RESULT 6

AX180276/c

LOCUS AX180276 132 bp DNA linear PAT 06-AUG-2001
DEFINITION Sequence 34 from Patent WO0146228.
ACCESSION AX180276
VERSION AX180276.1 GI:15132247
KEYWORDS
SOURCE
ORGANISM synthetic construct
synthetic construct
artificial sequences.

REFERENCE
AUTHORS Crowe, J.S. and Ellis, J.H.
TITLE Nucleic acid vaccination

JOURNAL

Patent: WO 0146228-A 34 28-JUN-2001;
GLAXO GROUP LIMITED (GB)

FEATURES

source

Location/Qualifiers
1. 132
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Oligonucleotide"

ORIGIN

Query Match 54.7%; Score 32.8; DB 6; Length 132;
Best Local Similarity 71.7%; Pred. No. 2.1;
Matches 43; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 1 GGAAGTACCGCTCCACCTGCACACGCGGTTCACAAAGCGCCGACACCTGCAGCTGGGCCA 60
|||||
Db 111 GGAGCTGCTGCTCCACCTGCACACGCGGTTCAGACACCTGCAGACACGCTCGGCGCCA 52

RESULT 7

AX180275

LOCUS AX180275 144 bp DNA linear PAT 06-AUG-2001
DEFINITION Sequence 33 from Patent WO0146228.
ACCESSION AX180275
VERSION AX180275.1 GI:15132246
KEYWORDS
SOURCE
ORGANISM synthetic construct
synthetic construct
artificial sequences.

REFERENCE
AUTHORS Crowe, J.S. and Ellis, J.H.
TITLE Nucleic acid vaccination
JOURNAL Patent: WO 0146228-A 33 28-JUN-2001;
GLAXO GROUP LIMITED (GB)

FEATURES

source

Location/Qualifiers
1. 144
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Oligonucleotide"

ORIGIN

Query Match 54.7%; Score 32.8; DB 6; Length 144;
Best Local Similarity 71.7%; Pred. No. 2.1;
Matches 43; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 1 GGAAGTACCGCTCCACCTGCACACGCGGTTCACAAAGCGCCGACACCTGCAGCTGGGCCA 60
|||||
Db 28 GGAGCTGCTGCTCCACCTGCACACGCGGTTCAGACACCTGCAGACACGCTCGGCGCCA 87

RESULT 8

AX703425

LOCUS AX703425 93 bp DNA linear PAT 03-APR-2003
DEFINITION Sequence 3 from Patent WO02086505.
ACCESSION AX703425
VERSION AX703425.1 GI:29538416
KEYWORDS
SOURCE
ORGANISM synthetic construct
synthetic construct
artificial sequences.

REFERENCE
AUTHORS Benson, R.S.
TITLE Intracellular analysis
JOURNAL Patent: WO 02086505-A 3 31-OCT-2002;
THE VICTORIA UNIVERSITY OF MANCHESTER (GB)

FEATURES
source
Location/Qualifiers
1. 93
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Artificial epitope construct"

ORIGIN

Query Match 52.3%; Score 31.4; DB 6; Length 93;
Best Local Similarity 85.4%; Pred. No. 6.1;
Matches 35; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GGAAGTACCGCTCCACCTGCACACGGGGTCAACAAGCGGCC 41
DB 38 GGATCTACCGCTCTCTCTGCCACGAGTCAACAAGCGCAC 78

RESULT 9
AX703427 156 bp DNA linear PAT 03-APR-2003
LOCUS
DEFINITION Sequence 5 from Patent WO02086505.
ACCESSION AX703427
VERSION AX703427.1 GI:29538418
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source

1 Benson, R.S.
Intracellular analysis
Patent: WO 02086505-A 5 31-OCT-2002;
THE VICTORIA UNIVERSITY OF MANCHESTER (GB)
Location/Qualifiers
1. 156
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Artificial epitope construct"

Query Match 52.3%; Score 31.4; DB 6; Length 156;
Best Local Similarity 85.4%; Pred. No. 6.1;
Matches 35; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GGAAGTACCGCTCCACCTGCACACGGGGTCAACAAGCGGCC 41
DB 38 GGATCTACCGCTCTCTCTGCCACGAGTCAACAAGCGCAC 78

RESULT 10
AX703428 157 bp DNA linear PAT 03-APR-2003
LOCUS
DEFINITION Sequence 6 from Patent WO02086505.
ACCESSION AX703428
VERSION AX703428.1 GI:29538419
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source

1 Benson, R.S.
Intracellular analysis
Patent: WO 02086505-A 6 31-OCT-2002;
THE VICTORIA UNIVERSITY OF MANCHESTER (GB)
Location/Qualifiers
1. 157
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Artificial epitope construct"

Query Match 52.3%; Score 31.4; DB 6; Length 157;
Best Local Similarity 85.4%; Pred. No. 6.1;
Matches 35; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GGAAGTACCGCTCCACCTGCACACGGGGTCAACAAGCGGCC 41
DB 38 GGATCTACCGCTCTCTCTGCCACGAGTCAACAAGCGCAC 78

RESULT 11
AX703431 157 bp DNA linear PAT 03-APR-2003
LOCUS
DEFINITION Sequence 9 from Patent WO02086505.
ACCESSION AX703431
VERSION AX703431.1 GI:29538422
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source

1 Benson, R.S.
Intracellular analysis
Patent: WO 02086505-A 9 31-OCT-2002;
THE VICTORIA UNIVERSITY OF MANCHESTER (GB)
Location/Qualifiers
1. 157
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Artificial epitope construct"

Query Match 52.3%; Score 31.4; DB 6; Length 157;
Best Local Similarity 85.4%; Pred. No. 6.1;
Matches 35; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GGAAGTACCGCTCCACCTGCACACGGGGTCAACAAGCGGCC 41
DB 38 GGATCTACCGCTCTCTCTGCCACGAGTCAACAAGCGCAC 78

RESULT 12
AX703433 3621 bp DNA linear PAT 03-APR-2003
LOCUS
DEFINITION Sequence 11 from Patent WO02086505.
ACCESSION AX703433
VERSION AX703433.1 GI:29538424
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source

1 Benson, R.S.
Intracellular analysis
Patent: WO 02086505-A 11 31-OCT-2002;
THE VICTORIA UNIVERSITY OF MANCHESTER (GB)
Location/Qualifiers
1. 3621
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Engineered construct"

Query Match 52.3%; Score 31.4; DB 6; Length 3621;
Best Local Similarity 85.4%; Pred. No. 4.3;
Matches 35; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GGAAGTACCGCTCCACCTGCACACGGGGTCAACAAGCGGCC 41
DB 659 GGATCTACCGCTCTCTCTGCCACGAGTCAACAAGCGCAC 699

RESULT 13
AX703435 6877 bp DNA linear PAT 03-APR-2003
LOCUS
DEFINITION Sequence 13 from Patent WO02086505.
ACCESSION AX703435
VERSION AX703435.1 GI:29538426
KEYWORDS
SOURCE
synthetic construct

ORGANISM synthetic construct
artificial sequences.

REFERENCE 1
AUTHORS Benson, R.S.
TITLE Intracellular analysis
JOURNAL Patent: WO 02086505-A 13 31-OCT-2002;
THE VICTORIA UNIVERSITY OF MANCHESTER (GB)

FEATURES
source 1.6877
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Engineered construct"

ORIGIN

Query Match 52.3%; Score 31.4; DB 6; Length 6877;
Best Local Similarity 85.4%; Pred. No. 4;
Matches 35; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 GGAAGTACCGCTCCACCTGCACACGCGGCTCACAAGCGCGCC 41
Db 701 GGATCTACCGCTCTCTCTCCACGAGTCAACGCGCACCC 741

RESULT 14
AX180282/c 132 bp DNA PAT 06-AUG-2001

LOCUS AX180282
DEFINITION Sequence 40 from Patent WO0146228.
ACCESSION AX180282
VERSION AX180282.1 GI:15132253

KEYWORDS

SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.

REFERENCE 1
AUTHORS Crowe, J.S. and Ellis, J.H.
TITLE Nucleic acid vaccination
JOURNAL Patent: WO 0146228-A 40 28-JUN-2001;
GLAXO GROUP LIMITED (GB)

FEATURES
source 1.132
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Oligonucleotide"

ORIGIN

Query Match 52.0%; Score 31.2; DB 6; Length 132;
Best Local Similarity 70.0%; Pred. No. 7.2;
Matches 42; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy 1 GGAAGTACCGCTCCACCTGCACACGCGGCTCACAAGCGCGCCAGACACTGACCTGGCGCA 60
Db 111 GGAGCTGTGCTCCACCTGCACACGCGGCTGTTGACGACACTGACAACCGTCCGGGCGCA 52

RESULT 15
AX180281 144 bp DNA PAT 06-AUG-2001

LOCUS AX180281
DEFINITION Sequence 39 from Patent WO0146228.
ACCESSION AX180281
VERSION AX180281.1 GI:15132252

KEYWORDS

SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.

REFERENCE 1
AUTHORS Crowe, J.S. and Ellis, J.H.
TITLE Nucleic acid vaccination
JOURNAL Patent: WO 0146228-A 39 28-JUN-2001;
GLAXO GROUP LIMITED (GB)

FEATURES
source 1.144
Location/Qualifiers

/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Oligonucleotide"

ORIGIN

Query Match 52.0%; Score 31.2; DB 6; Length 144;
Best Local Similarity 70.0%; Pred. No. 7.2;
Matches 42; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy 1 GGAAGTACCGCTCCACCTGCACACGCGGCTCACAAGCGCGCCAGACACTGACCTGGCGCA 60
Db 28 GGAGCTGTGCTCCACCTGCACACGCGGCTGTTGACGACACTGACAACCGTCCGGGCGCA 87

Search completed: January 15, 2005, 22:36:53
Job time : 722.8 secs

This Page Blank (uspto)

```

Query Match      100.0%; Score 60; DB 14; Length 60;
Best Local Similarity 100.0%; Pred. No. 1.4e-11;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps
OY      1 GGCAGACCCGACCCGCGACACGGGGGTCAACAAGCCGCGACACTGACCTGGCCA 60
|||||

```

Db 1 GGCAGACCGCACCGCCCGCACACGGGGTTCACAGCGCGCCAGACACTCGACCTGGCGCA 60

RESULT 2

US-10-057-136-9
; Sequence 9, Application US/10057136
; Publication No. US20030021770A1
; GENERAL INFORMATION:
; APPLICANT: SCHLON, JEFFREY
; APPLICANT: KANTOR, JUDITH
; APPLICANT: KUFEL, DONALD
; APPLICANT: PANICALI, DENNIS
; APPLICANT: GRITZ, LINDA
; TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
; FILE REFERENCE: 700953/47113C
; CURRENT APPLICATION NUMBER: US/10/057,136
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/366,670
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: PCT/US98/03693
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/038,253
; PRIOR FILING DATE: 1997-02-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-136-9

Query Match 86.7%; Score 52; DB 14; Length 60;
Best Local Similarity 91.7%; Pred. No. 8,9e-09;

Matches 55; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 GGCAGACCGCACCGCCCGCACACGGGGTTCACAGCGCGCCAGACACTCGACCTGGCGCA 60
Db 1 GGAAGTACCGCTCCACCTGCACACGGGGTTCACAAGCGCGCCAGACACTCGACCTGGCGCA 60

RESULT 3

US-10-406-317-41
; Sequence 41, Application US/10406317
; Publication No. US20040019195A1
; GENERAL INFORMATION:
; APPLICANT: SCHLON, JEFFREY
; APPLICANT: HODGE, JAMES
; APPLICANT: PANICALI, DENNIS
; TITLE OF INVENTION: A recombinant vector expressing multiple constitutively
; FILE REFERENCE: 38163-0189
; CURRENT APPLICATION NUMBER: US/10/406,317
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: US/09/856,988
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: PCT/US99/26866
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/111,582
; PRIOR FILING DATE: 1998-12-09
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41
; LENGTH: 2297
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: VECTOR
US-10-406-317-41

Query Match 86.7%; Score 52; DB 16; Length 2297;
Best Local Similarity 91.7%; Pred. No. 5,8e-09;

Matches 55; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 GGCAGACCGCACCGCCCGCACACGGGGTTCACAGCGCGCCAGACACTCGACCTGGCGCA 60
Db 706 GGAAGTACCGCTCCACCTGCACACGGGGTTCACAAGCGCGCCAGACACTCGACCTGGCGCA 765

RESULT 4

US-10-057-136-12
; Sequence 12, Application US/10057136
; Publication No. US20030021770A1
; GENERAL INFORMATION:
; APPLICANT: SCHLON, JEFFREY
; APPLICANT: KANTOR, JUDITH
; APPLICANT: KUFEL, DONALD
; APPLICANT: PANICALI, DENNIS
; APPLICANT: GRITZ, LINDA
; TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
; FILE REFERENCE: 700953/47113C
; CURRENT APPLICATION NUMBER: US/10/057,136
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/366,670
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: PCT/US98/03693
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/038,253
; PRIOR FILING DATE: 1997-02-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-136-12

Query Match 61.0%; Score 36.6; DB 14; Length 60;
Best Local Similarity 76.3%; Pred. No. 0.0021;

Matches 45; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 1 GGCAGACCGCACCGCCCGCACACGGGGTTCACAGCGCGCCAGACACTCGACCTGGCGCC 59
Db 1 GGTAGTACAGCGCGCACCGCCGACATGGCGTTCAGAGCCCTCGGATTCAGAGCCGGCGCC 59

RESULT 5

US-10-057-136-4
; Sequence 4, Application US/10057136
; Publication No. US20030021770A1
; GENERAL INFORMATION:
; APPLICANT: SCHLON, JEFFREY
; APPLICANT: KANTOR, JUDITH
; APPLICANT: KUFEL, DONALD
; APPLICANT: PANICALI, DENNIS
; APPLICANT: GRITZ, LINDA
; TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
; FILE REFERENCE: 700953/47113C
; CURRENT APPLICATION NUMBER: US/10/057,136
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/366,670
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: PCT/US98/03693
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/038,253
; PRIOR FILING DATE: 1997-02-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-136-4

```

; NUMBER OF S
; SOFTWARE: F
; SEQ ID NO 11
; LENGTH: 60
;

```

; PRIOR FILING DATE: 1998-
 ; PRIOR FILING DATE: 1997-
 ; NUMBER OF SEQ ID NOS: 20

```
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 13
LENGTH: 78
TYPE: DNA
ORGANISM: Homo sapiens
US-10-057-136-13
```

```
Query Match
Best Local Similarity 50.3%; Score 30.2; DB 14; Length 78;
Pred. No. 0.35;
Matches 41; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
```

```
QY 1 GGCAGACCCGACCGCCCGCCGACACGGGGTTCACAGCGCCGACAGACTGACCTGCGCC 59
DB 1 GGCTCCACCGACCGCCCGCCGACCGGTGTCACTCGGCCCGGACACAGGCGGCGCC 59
```

```
RESULT 10
US-10-447-839A-75/C
Sequence 75, Application US/10447839A
Publication No. US20040018181A1
GENERAL INFORMATION:
APPLICANT: Kufe, Donald W.
APPLICANT: Kharbada, Suresnder
APPLICANT: Weitman, Steven D.
TITLE OF INVENTION: MUC1 INTERFERENCE RNA COMPOSITIONS AND METHODS DERIVED THEREFROM
FILE REFERENCE: 1000.05.009
CURRENT APPLICATION NUMBER: US/10/447,839A
PRIORITY FILING DATE: 2003-05-29
PRIOR APPLICATION NUMBER: 10/293,391
PRIOR FILING DATE: 2002-11-13
PRIOR APPLICATION NUMBER: 09/951,938
PRIOR FILING DATE: 2001-09-11
PRIOR APPLICATION NUMBER: 60/231,841
NUMBER OF SEQ ID NOS: 109
SOFTWARE: PatentIn version 3.2
SEQ ID NO 75
LENGTH: 1424
TYPE: RNA
ORGANISM: ARTIFICIAL
FEATURE:
OTHER INFORMATION: Synthesized Sequence
US-10-447-839A-75
```

```
Query Match
Best Local Similarity 50.3%; Score 30.2; DB 16; Length 1424;
Pred. No. 0.25;
Matches 41; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
```

```
QY 1 GGCAGACCCGACCGCCCGCCGACACGGGGTTCACAGCGCCGACAGACTGACCTGCGCC 59
DB 1041 GGCTCCACCGCCCGCCCGCCGACCGGTGTCACTCGGCCCGGACACAGGCGGCGCC 983
```

```
RESULT 11
US-10-447-839A-20
Sequence 20, Application US/10447839A
Publication No. US20040018181A1
GENERAL INFORMATION:
APPLICANT: Kufe, Donald W.
APPLICANT: Kharbada, Suresnder
APPLICANT: Weitman, Steven D.
TITLE OF INVENTION: MUC1 INTERFERENCE RNA COMPOSITIONS AND METHODS DERIVED THEREFROM
FILE REFERENCE: 1000.05.009
CURRENT APPLICATION NUMBER: US/10/447,839A
PRIORITY FILING DATE: 2003-05-29
PRIOR APPLICATION NUMBER: 10/293,391
PRIOR FILING DATE: 2002-11-13
PRIOR APPLICATION NUMBER: 09/951,938
PRIOR FILING DATE: 2001-09-11
PRIOR APPLICATION NUMBER: 60/231,841
NUMBER OF SEQ ID NOS: 109
SOFTWARE: PatentIn version 3.2
```

```
SEQ ID NO 20
LENGTH: 1428
TYPE: RNA
ORGANISM: RNA
US-10-447-839A-20
```

```
Query Match
Best Local Similarity 50.3%; Score 30.2; DB 16; Length 1428;
Pred. No. 0.25;
Matches 40; Conservative 1; Mismatches 18; Indels 0; Gaps 0;
```

```
QY 1 GGCAGACCCGACCGCCCGCCGACACGGGGTTCACAGCGCCGACAGACTGACCTGCGCC 59
DB 385 GGCTCCACCGCCCGCCCGCCGACCGGTGTCACTCGGCCCGGACACAGGCGGCGCC 443
```

```
RESULT 12
US-10-057-136-19
Sequence 19, Application US/10057136
Publication No. US2003002170A1
GENERAL INFORMATION:
APPLICANT: SCHLOM, JEFFREY
APPLICANT: KANTOR, JUDITH
APPLICANT: Kufe, Donald W.
APPLICANT: PANTCALI, DENNIS
APPLICANT: GRITZ, LINDA
TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
FILE REFERENCE: 700953/47113C
CURRENT APPLICATION NUMBER: US/10/057,136
PRIORITY FILING DATE: 2002-01-25
PRIOR APPLICATION NUMBER: 09/366,670
PRIOR FILING DATE: 1999-08-03
PRIOR APPLICATION NUMBER: PCT/US98/03693
PRIOR FILING DATE: 1998-02-24
PRIOR APPLICATION NUMBER: 60/038,253
PRIOR FILING DATE: 1997-02-24
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 19
LENGTH: 1527
TYPE: DNA
ORGANISM: Homo sapiens
US-10-057-136-19
```

```
Query Match
Best Local Similarity 50.3%; Score 30.2; DB 14; Length 1527;
Pred. No. 0.25;
Matches 41; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
```

```
QY 1 GGCAGACCCGACCGCCCGCCGACACGGGGTTCACAGCGCCGACAGACTGACCTGCGCC 59
DB 226 GGCTCCACCGCCCGCCCGCCGACCGGTGTCACTCGGCCCGGACACAGGCGGCGCC 284
```

```
RESULT 13
US-10-447-839A-19
Sequence 19, Application US/10447839A
Publication No. US20040018181A1
GENERAL INFORMATION:
APPLICANT: Kufe, Donald W.
APPLICANT: Kharbada, Suresnder
APPLICANT: Weitman, Steven D.
TITLE OF INVENTION: MUC1 INTERFERENCE RNA COMPOSITIONS AND METHODS DERIVED THEREFROM
FILE REFERENCE: 1000.05.009
CURRENT APPLICATION NUMBER: US/10/447,839A
PRIORITY FILING DATE: 2003-05-29
PRIOR APPLICATION NUMBER: 10/293,391
PRIOR FILING DATE: 2002-11-13
PRIOR APPLICATION NUMBER: 09/951,938
PRIOR FILING DATE: 2001-09-11
PRIOR APPLICATION NUMBER: 60/231,841
NUMBER OF SEQ ID NOS: 109
SOFTWARE: PatentIn version 3.2
```

SEQ ID NO 19
LENGTH: 1799
TYPE: RNA
ORGANISM: RNA
US-10-447-839A-19

Query Match 50.3%; Score 30.2; DB 16; Length 1799;
Best Local Similarity 67.8%; Pred. No. 0.24;
Matches 40; Conservative 1; Mismatches 18; Indels 0; Gaps 0;

QY 1 GGCAGACCGCACCGCCCGCACAGGGGTTCACAAGCGCCGACACTGACCTGCGCC 59
DB 457 GGCTCCACCGCCCCCGCACCGCGGTGTCACTCGGCCCCGAGACACAGCGCGCCCC 511

RESULT 14
US-09-964-824A-573
Sequence 573, Application US/09964824A
Patent No. US20020102531A1
GENERAL INFORMATION:

APPLICANT: Horrigan, Stephen
TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
FILE REFERENCE: 689290-73
CURRENT APPLICATION NUMBER: US/09/964,824A
CURRENT FILING DATE: 2001-09-27
PRIOR APPLICATION NUMBER: US/60/236,033
PRIOR FILING DATE: 2000-09-28
PRIOR APPLICATION NUMBER: US/60/236,032
PRIOR FILING DATE: 2000-09-28
PRIOR APPLICATION NUMBER: US/60/236,028
PRIOR FILING DATE: 2000-09-28
NUMBER OF SEQ ID NOS: 583
SOFTWARE: PatentIn version 3.0
SEQ ID NO 573
LENGTH: 1804
TYPE: DNA
ORGANISM: Homo sapiens
US-09-964-824A-573

Query Match 50.3%; Score 30.2; DB 9; Length 1804;
Best Local Similarity 69.5%; Pred. No. 0.24;
Matches 41; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1 GGCAGACCGCACCGCCCGCACAGGGGTTCACAAGCGCCGACACTGACCTGCGCC 59
DB 457 GGCTCCACCGCCCCCGCACCGCGGTGTCACTCGGCCCCGAGACACAGCGCGCCCC 515

RESULT 15

US-10-029-517-17
Sequence 17, Application US/10029517
Publication No. US20030148969A1
GENERAL INFORMATION:
APPLICANT: Susan J. Wiers
APPLICANT: Kenneth W. Doble
TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
FILE REFERENCE: RTS-0352
CURRENT APPLICATION NUMBER: US/10/029,517
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 107
SEQ ID NO 17
LENGTH: 1804
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (73)...(1500)
US-10-029-517-17

Query Match 50.3%; Score 30.2; DB 15; Length 1804;
Best Local Similarity 69.5%; Pred. No. 0.24;
Matches 41; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1 GGCAGACCGCACCGCCCGCACAGGGGTTCACAAGCGCCGACACTGACCTGCGCC 59
DB 457 GGCTCCACCGCCCCCGCACCGCGGTGTCACTCGGCCCCGAGACACAGCGCGCCCC 515

Search completed: January 16, 2005, 09:30:25
Job time : 183.4 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2005, 19:50:13 : Search time 35.2 Seconds
(without alignments)
1211.572 Million cell updates/sec

Title: US-10-057-136-8

Perfect score: 60
Sequence: 1 GGCAGCAGCCGACGCGCCGCGC.....CAGACACGACCTGCGCCA 60

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents NA:*
1: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/prodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/prodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/prodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/prodata/1/ina/6CTUS_COMB.seq:*
6: /cgn2_6/prodata/1/ina/Backfillseq1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	30.2	50.3	1804	US-10-029-517-17	Sequence 17, App1
2	30.2	50.3	8186	US-10-029-517-19	Sequence 19, App1
3	28.8	48.0	518	US-10-029-517-101	Sequence 101, App
4	28.6	47.7	572	US-10-029-517-18	Sequence 18, App1
5	28.6	47.7	1721	US-10-029-517-3	Sequence 3, App1
6	27.4	45.7	60	US-09-475-947A-246	Sequence 246, App
7	27.2	45.3	3343	US-10-029-517-102	Sequence 102, App
8	26	43.3	6192	US-08-479-537A-1	Sequence 1, App1
9	26	43.3	6192	US-09-083-116-1	Sequence 1, App1
10	26	43.3	6192	US-09-134-916A-1	Sequence 1, App1
11	26	43.3	6449	US-08-479-537A-4	Sequence 4, App1
12	26	43.3	6449	US-09-083-116-4	Sequence 4, App1
13	26	43.3	6449	US-09-134-916A-4	Sequence 4, App1
14	25.6	42.7	981	US-10-029-517-16	Sequence 16, App1
15	25.6	42.7	4041	US-09-569-611C-1	Sequence 1, App1
16	24.8	41.3	1174	US-09-034-985-1	Sequence 1, App1
17	24.8	41.3	1417	US-09-535-008-33	Sequence 33, App1
18	24.4	40.7	846	US-09-455-960-5	Sequence 5, App1
19	24.4	40.7	846	US-10-051-325-5	Sequence 5, App1
20	24.4	40.7	2500	US-09-455-960-2	Sequence 2, App1
21	24.4	40.7	2500	US-10-051-325-2	Sequence 2, App1
22	24.4	40.7	2748	US-09-799-451-598	Sequence 598, App
23	24.2	40.3	1392	US-09-489-039A-5663	Sequence 5663, App
24	24	40.0	2900	US-08-034-650-9	Sequence 9, App1
25	23.8	39.7	633	US-08-449-015-9	Sequence 9, App1
26	23.8	39.7	633	US-09-252-991A-3966	Sequence 3966, App
27	23.8	39.7	1899	US-09-252-991A-4034	Sequence 4034, App

28	23.8	39.7	1920	US-09-252-991A-4000	Sequence 4000, App
29	23.4	39.0	154	US-09-513-999C-26161	Sequence 26161, A
30	23.4	39.0	888	US-09-311-021-155	Sequence 155, App
31	23.4	39.0	920	US-09-620-312D-132	Sequence 132, App
32	23.4	39.0	4403765	US-09-103-840A-2	Sequence 2, App1
33	23.4	39.0	4411529	US-09-103-840A-1	Sequence 1, App1
34	23.2	38.7	3957	US-07-689-008-5	Sequence 5, App1
35	23.2	38.7	9940	US-07-689-008-1	Sequence 1, App1
36	23.2	38.3	274	US-09-621-976-16450	Sequence 16450, A
37	23	38.3	425	US-08-956-171E-3693	Sequence 3693, App
38	23	38.3	425	US-08-781-986A-3693	Sequence 3693, App
39	23	38.3	918	US-09-252-991A-9303	Sequence 9303, App
40	23	38.3	1428	US-09-252-991A-9266	Sequence 9266, App
41	23	38.3	1545	US-09-252-991A-9274	Sequence 9274, App
42	22.8	38.0	355	US-09-370-838-135	Sequence 135, App
43	22.8	38.0	355	US-09-854-133-135	Sequence 135, App
44	22.8	38.0	507	US-08-956-171E-3895	Sequence 3895, App
45	22.8	38.0	507	US-08-781-986A-3895	Sequence 3895, App

ALIGNMENTS

```

RESULT 1
US-10-029-517-17
; Sequence 17, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 17
; LENGTH: 1804
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (73)...(1500)
US-10-029-517-17

Query Match          50.3%; Score 30.2; DB 4; Length 1804;
Best Local Similarity 69.5%; Pred. No. 0.26;
Matches 41; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY      1 GGCAGCAGCCGACCGCGCGGCGGTCACAGGCGCGCAGACACTGACCTGGCC 59
Db      457 GGCCTCACCAGCCCGCCCGACCGACCGTGTACCTCGCGCCCGACACAGCGCCGCC 515

RESULT 2
US-10-029-517-19
; Sequence 19, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 19
; LENGTH: 8186
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 6899
OTHER INFORMATION: unknown

```

```
/ NAME/KEY: unsure
/ LOCATION: 7155
/ OTHER INFORMATION: unknown
/ NAME/KEY: unsure
/ LOCATION: 7184
/ OTHER INFORMATION: unknown
/ NAME/KEY: unsure
/ LOCATION: 7957
/ OTHER INFORMATION: unknown
/ NAME/KEY: intron
/ LOCATION: (2997)...(3498)
/ OTHER INFORMATION: intron 1
/ NAME/KEY: intron:exon junction
/ LOCATION: (3498)...(3499)
/ OTHER INFORMATION: intron 1:exon 2
/ NAME/KEY: exon
/ LOCATION: (3508)...(3599)
/ OTHER INFORMATION: exon 2d
/ NAME/KEY: exon:intron junction
/ LOCATION: (3982)...(3983)
/ OTHER INFORMATION: exon 2a:intron 2a
/ NAME/KEY: intron:exon junction
/ LOCATION: (4205)...(4206)
/ OTHER INFORMATION: intron 2c:exon 3c
/ NAME/KEY: intron:exon junction
/ LOCATION: (4259)...(4260)
/ OTHER INFORMATION: intron 2d:exon 3d
/ NAME/KEY: exon
/ LOCATION: (4260)...(4328)
/ OTHER INFORMATION: exon 3d
/ NAME/KEY: intron:exon junction
/ LOCATION: (4632)...(4633)
/ OTHER INFORMATION: intron 3:exon 4
/ NAME/KEY: exon
/ LOCATION: (4914)...(5035)
/ OTHER INFORMATION: exon 5
/ NAME/KEY: intron
/ LOCATION: (5266)...(6293)
/ OTHER INFORMATION: intron 6
US-10-029-517-19
```

```
Query Match
Best Local Similarity 50.3%; Score 30.2; DB 4; Length 8186;
Matches 41; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
```

```
QY 1 GGCAGCAGCCGACCGCCCGGACAGCGGGTCAAGCGCCGACACTGACCTGGCGCC 59
DB 3825 GGCTCCACCGCCCGCCCGGACGCGGTCACTCGGCGCCCGGACAGCGCCGCGCC 3883
```

RESULT 3

```
US-10-029-517-101
/ Sequence 101, Application US/10029517
/ Patent No. 6716627
/ GENERAL INFORMATION:
/ APPLICANT: Kenneth W. Dobie
/ TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
/ FILE REFERENCE: RTS-0352
/ CURRENT APPLICATION NUMBER: US/10/029,517
/ NUMBER OF SEQ ID NOS: 107
/ SEQ ID NO 101
/ LENGTH: 518
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
US-10-029-517-101
```

```
Query Match
Best Local Similarity 48.0%; Score 28.8; DB 4; Length 518;
Matches 39; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
```

```
QY 1 GGCAGCAGCCGACCGCCCGGACAGCGGGTCAAGCGCCGACACTGACCTGCGCC 56
DB 462 GGCTCCACCGCCCGCCCGGACGCGGTCACTCGGCGCCCGGACAGCGCGCGCC 517
```

RESULT 4

```
US-10-029-517-18
/ Sequence 18, Application US/10029517
/ Patent No. 6716627
/ GENERAL INFORMATION:
/ APPLICANT: Kenneth W. Dobie
/ TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
/ FILE REFERENCE: RTS-0352
/ CURRENT APPLICATION NUMBER: US/10/029,517
/ NUMBER OF SEQ ID NOS: 107
/ SEQ ID NO 18
/ LENGTH: 572
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (67)...(572)
US-10-029-517-18
```

```
Query Match
Best Local Similarity 47.7%; Score 28.6; DB 4; Length 572;
Matches 40; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
```

```
QY 1 GGCAGCAGCCGACCGCCCGGACAGCGGGTCAAGCGCCGACACTGACCTGGCGCC 59
DB 478 GGCTCCACCGCCCGCCCGGACGCGGTCACTCGGCGCCCGGACAGCGCGCGCC 536
```

RESULT 5

```
US-10-029-517-3
/ Sequence 3, Application US/10029517
/ Patent No. 6716627
/ GENERAL INFORMATION:
/ APPLICANT: Kenneth W. Dobie
/ TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
/ FILE REFERENCE: RTS-0352
/ CURRENT APPLICATION NUMBER: US/10/029,517
/ NUMBER OF SEQ ID NOS: 107
/ SEQ ID NO 3
/ LENGTH: 1721
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (58)...(1605)
US-10-029-517-3
```

```
Query Match
Best Local Similarity 47.7%; Score 28.6; DB 4; Length 1721;
Matches 40; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
```

```
QY 1 GGCAGCAGCCGACCGCCCGGACAGCGGGTCAAGCGCCGACACTGACCTGGCGCC 59
DB 442 GGCTCCACCGCCCGCCCGGACGCGGTCACTCGGCGCCCGGACAGCGCGCGCC 500
```

RESULT 6

```
US-09-475-947A-246
/ Sequence 246, Application US/09475947A
/ Patent No. 6472154
/ GENERAL INFORMATION:
/ APPLICANT: Garner, Harold R.
/ APPLICANT: Wren, Jonathan D.
/ APPLICANT: Minna, John D.
```


US-09-083-116-1
Sequence 1, Application US/09083116
Patent No. 6203795
GENERAL INFORMATION:
APPLICANT: CHAMON, Pierre
APPLICANT: KIENY, Marie-Paule
APPLICANT: LATHE, Richard
APPLICANT: HAREUVENI, Mara
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
TREATMENT OR PREVENTION OF A MALIGNANT TUMOR
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESSES:
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/083,116
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/479,537
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/FR91/00835
FILING DATE: 23-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/039,320
FILING DATE: 04-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/403,576
FILING DATE: 14-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Teskin, Robin L.
REGISTRATION NUMBER: 35,030
REFERENCE/DOCKET NUMBER: 017753-025
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 6192 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 56..120
FEATURE:
NAME/KEY: repeat_region
LOCATION: 439..5239
OTHER INFORMATION: /note= "The nucleotides spanning
OTHER INFORMATION: 439-5239 constitute a repeated region wherein the repeat is 6
OTHER INFORMATION: nucleotides and encodes 20 amino acids, 17 of which are fixed
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 121..6166
FEATURE:
NAME/KEY: repeat_region
LOCATION: 457
OTHER INFORMATION: /note= "Nucleotide 457 is X1 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
OTHER INFORMATION: or CCG, and Ala = GCT, GCC, GCA, or GCG."

NAME/KEY: repeat_region
LOCATION: 487
OTHER INFORMATION: /note= "Nucleotide 487 is Y = NNN
OTHER INFORMATION: which is the codon for Thr or Asn wherein Thr = ACT, ACC, ACA
OTHER INFORMATION: or ACG, and Asn = AAT or AAC."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 496
OTHER INFORMATION: /note= "Nucleotide 496 is X2 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
OTHER INFORMATION: or CCG, and Ala = GCT, GCC, GCA, or GCG."
US-09-083-116-1
Query Match 43.3%; Score 26; DB 3; Length 6192;
Best Local Similarity 59.3%; Pred. No. 7;
Matches 35; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
QY 1 GGCAGCAGCCGACCGCCGACACGAGGCTCAAGCGGCGACACTCGACTCGCGCC 59
DB 442 GGCTCCACCGCCCGCCNNNGCCCAAGGTGTCTACTCGGCGCCCGGACNNAGCCGNNNCC 500
RESULT 10
US-09-134-916A-1
Sequence 1, Application US/09134916A
Patent No. 6328956
GENERAL INFORMATION:
APPLICANT: CHAMON, Pierre
APPLICANT: KIENY, Marie-Paule
APPLICANT: LATHE, Richard
APPLICANT: HAREUVENI, Mara
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
TREATMENT OR PREVENTION OF A MALIGNANT TUMOR
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESSES:
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/134,916A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/479,537
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: FR 90/13101
FILING DATE: 23-OCT-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/FR91/00835
FILING DATE: 23-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/039,320
FILING DATE: 04-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/403,576
FILING DATE: 14-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Teskin, Robin L.
REGISTRATION NUMBER: 35,030
REFERENCE/DOCKET NUMBER: 017753-025
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:

LENGTH: 6192 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 58..120
FEATURE:
NAME/KEY: repeat_region
LOCATION: 439..5239
OTHER INFORMATION: /note= "The nucleotides spanning 439-5239 constitute a repeated region wherein the repeat is 6
OTHER INFORMATION: 439-5239 constitute a repeated region wherein the repeat is 6
OTHER INFORMATION: nucleotides and encodes 20 amino acids, 17 of which are fixed
OTHER INFORMATION: The number of such repeats varies from 1 to 80."
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 121..6166
FEATURE:
NAME/KEY: repeat_region
LOCATION: 457
OTHER INFORMATION: /note= "Nucleotide 457 is X1 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
OTHER INFORMATION: or CCG, and Ala = GCT, GCC, GCA, or GCG."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 487
OTHER INFORMATION: /note= "Nucleotide 487 is Y = NNN
OTHER INFORMATION: which is the codon for Thr or Asn wherein Thr = ACT, ACC, ACA
OTHER INFORMATION: or ACG, and Asn = AAT or AAC."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 496
OTHER INFORMATION: /note= "Nucleotide 496 is X2 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
OTHER INFORMATION: or CCG, and Ala = GCT, GCC, GCA, or GCG."
US-09-134-916A-1

Query Match 43.3%; Score 26; DB 3; Length 6192;
Best Local Similarity 59.3%; Pred. No. 7;
Matches 35; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 1 GGCAGACCGCAGCCGCCGCGGTCACACAGCGCGCCGACACTGAGCTGGCC 59
DB 442 GGCCTCACCGCCGCCCCNNNGCCACGCGTGTACCTCGGCCCCGACNNNAGCCGNNCC 500

RESULT 11
US-08-479-537A-4
Sequence 4, Application US/08479537A
Patent No. 5861381
GENERAL INFORMATION:
APPLICANT: CHAMBER, Pierre
APPLICANT: KIENY, Marie-Paule
APPLICANT: LATHE, Richard
APPLICANT: HAREVENT, Maira
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
TITLE OF INVENTION: TREATMENT OR PREVENTION OF A MALIGNANT TUMOR
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSER: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/479,537A

FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 90/13101
FILING DATE: 23-OCT-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/FR91/00835
FILING DATE: 23-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/039,320
FILING DATE: 04-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/403,576
FILING DATE: 14-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Teskin, Robin L.
REGISTRATION NUMBER: 35,030
REFERENCE/DOCKET NUMBER: 017753-025
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 6449 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 58..120
FEATURE:
NAME/KEY: repeat_region
LOCATION: 439..5239
OTHER INFORMATION: /note= "The nucleotides spanning 439-5239 constitute a repeated region wherein the repeat is
OTHER INFORMATION: 439-5239 constitute a repeated region wherein the repeat is
OTHER INFORMATION: nucleotides and encodes 20 amino acids, 17 of which are fixe
OTHER INFORMATION: The number of such repeats varies from 1 to 80."
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 121..5661
FEATURE:
NAME/KEY: repeat_region
LOCATION: 457
OTHER INFORMATION: /note= "Nucleotide 457 is X1 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
OTHER INFORMATION: or CCG, and Ala = GCT, GCC, GCA, or GCG."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 487
OTHER INFORMATION: /note= "Nucleotide 487 is Y = NNN
OTHER INFORMATION: which is the codon for Thr or Asn wherein Thr = ACT, ACC, ACA
OTHER INFORMATION: or ACG, and Asn = AAT or AAC."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 496
OTHER INFORMATION: /note= "Nucleotide 496 is X2 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
OTHER INFORMATION: or CCG, and Ala = GCT, GCC, GCA, or GCG."
US-08-479-537A-4

Query Match 43.3%; Score 26; DB 2; Length 6449;
Best Local Similarity 59.3%; Pred. No. 7;
Matches 35; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 1 GGCAGACCGCAGCCGCCGCGGTCACACAGCGCGCCGACACTGAGCTGGCC 59
DB 442 GGCCTCACCGCCGCCCCNNNGCCACGCGTGTACCTCGGCCCCGACNNNAGCCGNNCC 500

RESULT 12
US-09-083-116-4
Sequence 4, Application US/09083116

Patent No. 6203795

GENERAL INFORMATION:

APPLICANT: CHAMRON, Pierre

APPLICANT: KIENY, Marie-Paule

APPLICANT: LATHE, Richard

APPLICANT: HAREUVEN, Mara

TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE

TITLE OF INVENTION: TREATMENT OR PREVENTION OF A MALIGNANT TUMOR

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:

ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.

STREET: P.O. Box 1404

CITY: Alexandria

STATE: Virginia

COUNTRY: United States

ZIP: 2213-1404

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/083,116

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/479,537

FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: WO PCT/FR91/00835

FILING DATE: 23-OCT-1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/039,320

FILING DATE: 04-APR-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/403,576

FILING DATE: 14-MAR-1995

ATTORNEY/AGENT INFORMATION:

NAME: Teskin, Robin L.

REGISTRATION NUMBER: 35,030

REFERENCE/DOCKET NUMBER: 017753-025

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 836-6620

TELEFAX: (703) 836-2021

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 6449 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

FEATURE:

NAME/KEY: sig_peptide

LOCATION: 58..120

FEATURE:

NAME/KEY: repeat_region

LOCATION: 439..5239

OTHER INFORMATION: /note= "The nucleotides spanning

OTHER INFORMATION: 439-5239 constitute a repeated region wherein the repeat is 6

OTHER INFORMATION: nucleotides and encodes 20 amino acids, 17 of which are fixed

OTHER INFORMATION: The number of such repeats varies from 1 to 80."

FEATURE:

NAME/KEY: mat_peptide

LOCATION: 121..5661

FEATURE:

NAME/KEY: repeat_region

LOCATION: 457

OTHER INFORMATION: /note= "Nucleotide 457 is X1 = NNN

OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA

OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."

FEATURE:

NAME/KEY: repeat_region

LOCATION: 487

OTHER INFORMATION: /note= "Nucleotide 487 is Y = NNN

OTHER INFORMATION: which is the codon for Thr or Asn wherein Thr = ACT, ACC, ACA

OTHER INFORMATION: or ACG; and Asn = AAT or AAC."

FEATURE:

NAME/KEY: repeat_region

LOCATION: 496

OTHER INFORMATION: /note= "Nucleotide 496 is X2 = NNN

OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA

OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."

US-09-083-116-4

Query Match 43.3%; Score 26; DB 3; Length 6449;

Best Local Similarity 59.3%; Pred. No. 7;

Matches 35; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 1 GGCAGACCGGACCGCCCGGACGAGGTCAGACGCGGACGACCTCGACCTGCGCC 59

DB 442 GGCCTCACCGCCCGCCNNNGCCGACGCTGTCACCTCGGCCCGGACNNNAGCCGNNNCC 500

RESULT 13

US-09-134-916A-4

Sequence 4, Application US/09134916A

Patent No. 6328956

GENERAL INFORMATION:

APPLICANT: CHAMRON, Pierre

APPLICANT: KIENY, Marie-Paule

APPLICANT: LATHE, Richard

APPLICANT: HAREUVEN, Mara

TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE

TITLE OF INVENTION: TREATMENT OR PREVENTION OF A MALIGNANT TUMOR

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:

ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.

STREET: P.O. Box 1404

CITY: Alexandria

STATE: Virginia

COUNTRY: United States

ZIP: 2213-1404

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/134,916A

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/479,537

FILING DATE: 07-JUN-1995

APPLICATION NUMBER: FR 90/13101

FILING DATE: 23-OCT-1990

PRIOR APPLICATION DATA:

APPLICATION NUMBER: WO PCT/FR91/00835

FILING DATE: 23-OCT-1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/039,320

FILING DATE: 04-APR-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/403,576

FILING DATE: 14-MAR-1995

ATTORNEY/AGENT INFORMATION:

NAME: Teskin, Robin L.

REGISTRATION NUMBER: 35,030

REFERENCE/DOCKET NUMBER: 017753-025

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 836-6620

TELEFAX: (703) 836-2021

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 6449 base pairs

TYPE: nucleic acid

```

1 STRANDNESS: single
2 TOPOLOGY: linear
3 MOLECULE TYPE: DNA (genomic)
4 FEATURE:
5 NAME/KEY: sig_peptide
6 LOCATION: 58..120
7 FEATURE:
8 NAME/KEY: repeat_region
9 LOCATION: 439..5239
10 OTHER INFORMATION: /note= "The nucleotides spanning
11 OTHER INFORMATION: 439-5239 constitute a repeated region wherein the repeat is 6
12 OTHER INFORMATION: nucleotides and encodes a 20 amino acids, 17 of which are fixed
13 OTHER INFORMATION: The number of such repeats varies from 1 to 80."
14 FEATURE:
15 NAME/KEY: mac_peptide
16 LOCATION: 121..5661
17 FEATURE:
18 NAME/KEY: repeat_region
19 LOCATION: 457
20 OTHER INFORMATION: /note= "Nucleotide 457 is X1 = NNN
21 OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCG, CCA
22 OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
23 FEATURE:
24 NAME/KEY: repeat_region
25 LOCATION: 487
26 OTHER INFORMATION: /note= "Nucleotide 487 is Y = NNN
27 OTHER INFORMATION: which is the codon for Pro or Asn wherein Thr = ACT, ACC, ACA
28 OTHER INFORMATION: or ACG; and Asn = AAT or AAC."
29 FEATURE:
30 NAME/KEY: repeat_region
31 LOCATION: 496
32 OTHER INFORMATION: /note= "Nucleotide 496 is X2 = NNN
33 OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCG, CCA
34 OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
35
36 US-09-134-916A-4

```

```

Query Match          43.3%; Score 26; DB 3; Length 6449;
Best Local Similarity 59.3%; Pred. No. 7;
Matches 35; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

OY      1  GGCAGCAGCCGACAGCCGCCGACAGAGGGGTCAACAAGCGCCGACAGACTGCAGCC 59
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      442 GGCCTCCAGCCGCCCCCNNGGCCAGGTGTATCCTCGAGCCCGAGCCNNAAGCCGANNCC 500

RESULT 14
US-10-029-517-16
; Sequence 16, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Susan J. Myers
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RRS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 107
SEQ ID NO 16
LENGTH: 981
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: exon:exon junction
; LOCATION: (464)...(465)
; OTHER INFORMATION: exon 3b:exon 4
US-10-029-517-16

```

Query Match	42.7%	Score 25.6	DB 4	Length 991
Best Local Similarity	66.1%	Pred. No. 8.1		
Matches 37, Conservative	0	Mismatches 19	Indels 0	Gaps 0
0y	1	GGACGACCGGACCGCCGCGACGAGGGGTCACAGCGGCCGACGACTGCACCTGC 56		

Db 21 GGCTCAGCGCCCCCGCCGCCATGATGTACTCTCGGCGCCCCGAGCAAGAGCCCCG 76

RESULT 15
US-09-569-611C-1/c
; Sequence 1, Application US/09569611C
; Patent No. 6720182
; GENERAL INFORMATION:
; APPLICANT: SAVITZKY et al.
; TITLE OF INVENTION: ALTERNATIVE SPLICING VARIANTS
; FILE REFERENCE: 2786-0151P
; CURRENT APPLICATION NUMBER: US/09/569, 611C
; CURRENT FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4041
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-569-611C-1

	Query Match	Score	DB	Length
1	41.7%	25	4	4041
	Best Local Similarity	64.9%	Pred. No. 14	
Matches	37	Conservative	0	Mismatches 20; Indels 0; Gaps 0;

	Query	DB
1	GGCAGCACCGACCGCCCGCACACGGGGGTACAAGCGCGCCAGACACTTCGACTCGG	57
2244	GGCGCGCTAGACCGCTGCGCTCAGACGGGGACCTTGTCTCTCTGATTCGGAATTCG	2188

Search completed: January 16, 2005, 03:01:42
Job time : 40.2 BECS

This Page Blank (uspto)

TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 14, row: A column: 06
High quality sequence stop: 604.
Location/Qualifiers
1. 604

FEATURES
source
/organism="Homo sapiens"
/mol_type="rRNA"
/db_xref="taxon:9606"
/clone="S21SNUS20-14-A06"
/sex="F"
/issue_type="Stomach"
/cell_type="Floating aggregates"
/cell_line="SNU-520"
/lab_host="Top10F"
/clone_lib="S21SNUS20"
/note="Organ: Stomach; Vector: PT218RP1; Site 1: EcoRI;
Site 2: NotI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tobacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including EcoR
I site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized from oligo dT-selected mRNA by
priming with dT-tailed vector. The dT-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."

ORIGIN

Query Match 53.0%; Score 31.8; DB 4; Length 604;
Best Local Similarity 71.2%; Pred. No. 9.2;
Matches 42; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 1 GGCAGCAGCGCACCCGCCGACAGCGGGGTACAAAGCGCCGACACTGCGCTGGCC 59
Db 28 GGCTCAGCGCACCCGCCGACAGCGGTGTCACTCGCGCCCGGACACAGCGCGGCC 86

RESULT 4 BUS42454 877 bp mRNA linear EST 13-SEP-2002
BUS42454 AGENCOURT 10322173 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6574322
DEFINITION 5', mRNA sequence.
ACCESSION BUS42454
VERSION BUS42454.1 GI:22852937
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 877)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DTF
cDNA Library Preparation: Rubin Laboratory
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LNCM2769 row: h column: 02
High quality sequence stop: 760.
Location/Qualifiers
1. 877

FEATURES
source
/organism="Homo sapiens"
/mol_type="rRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6574322"
/issue_type="Carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 40"
/note="Organ: prostate; Vector: pORF7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(5). Library constructed by
ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH-MGC library."

ORIGIN

Query Match 53.0%; Score 31.8; DB 5; Length 877;
Best Local Similarity 71.2%; Pred. No. 9;
Matches 42; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 1 GGCAGCAGCGCACCCGCCGACAGCGGGGTACAAAGCGCCGACACTGCGCTGGCC 59
Db 215 GGCTCAGCGCACCCGCCGACAGCGGTGTCACTCGCGCCCGGACACAGCGCGGCC 273

RESULT 5 CA489836 959 bp mRNA linear EST 14-NOV-2002
CA489836 AGENCOURT 10810668 MAPcl Homo sapiens cDNA clone IMAGE:6722324 5',
DEFINITION mRNA sequence.
ACCESSION CA489836
VERSION CA489836.1 GI:24952627
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 959)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Kristi A. Eglund, Ira Pastan
cDNA Library Preparation: Invitrogen Corp
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LLM14284 row: n column: 20
High quality sequence start: 31
High quality sequence stop: 446.
Location/Qualifiers
1. 959

FEATURES
source

/organism="Homo sapiens"
/mol_type="rRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6722324"
/cell_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231,
HRT-EMBL, LNCap"
/lab_host="EMDH10B"
/clone_lib="MAPCL"
/note="Vector: PCMV-SPORT6; Site 1: EcoRV; Site 2: Not I;
Subtracted with brain, liver, lung, kidney and muscle.
Directionally cloned. Priming method: oligo-dT. Average

Insert size: 1800 bp. Library amplification: 26,000 fold.
 Kristi A. Egland, James J. Vincent, Robert Strausberg,
 Bungook Lee & Ira Pastan: Discovery of new breast
 cancer genes encoding membrane and secreted proteins.
 Manuscript submitted."

ORIGIN

Query Match 53.0%; Score 31.8; DB 6; Length 959;
 Best Local Similarity 71.2%; Pred. No. 9;
 Matches 42; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 1 GGCACACCGCCGACACCGCGGTCACAGCGCGACACTCGACCTGCGCC 59
 Db 469 GGCCTCCACCGACCCCGACCGGTCTCCTCGCGCCGACACCGCGCC 527

RESULT 6

AUI92332 499 bp mRNA linear EST 14-OCT-2003
 LOCUS AUI92332 Porphyra yezoensis TU-1 sporophytes Porphyra yezoensis
 DEFINITION CDNA clone PFL029d07_r 5', mRNA sequence.
 ACCESSION AUI92332
 VERSION AUI92332.1 GI:31930868
 KEYWORDS EST.
 SOURCE Porphyra yezoensis
 ORGANISM Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae;

REFERENCE

1 (bases 1 to 499)
 Asamizu,E., Nakajima,M., Kitade,Y., Saga,N., Nakamura,Y. and
 Tabata,S.
 TITLE COMPARISON OF RNA EXPRESSION PROFILES BETWEEN THE TWO GENERATIONS
 OF PORPHYRA YEZOENSIS (RHODOPHYTA), BASED ON EXPRESSED SEQUENCE TAG
 FREQUENCY ANALYSIS
 J. Phycol. 39 (5), 923-930 (2003)
 CONTACT: Erika Asamizu
 The First Laboratory for Plant Gene Research
 Kazusa DNA Research Institute
 Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
 Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES

1..499
 location/Qualifiers
 /organism="Porphyra yezoensis"
 /mol_type="mRNA"
 /strain="TU-1"
 /db_xref="taxon:2788"
 /clone="PFL029d07_r"
 /dev_stage="sporophytes"
 /clone_lib="Porphyra yezoensis TU-1 sporophytes"

JOURNAL

COMMENT The First Laboratory for Plant Gene Research
 Kazusa DNA Research Institute
 Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
 Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES

1..499
 location/Qualifiers
 /organism="Porphyra yezoensis"
 /mol_type="mRNA"
 /strain="TU-1"
 /db_xref="taxon:2788"
 /clone="PFL029d07_r"
 /dev_stage="sporophytes"
 /clone_lib="Porphyra yezoensis TU-1 sporophytes"

ORIGIN

Query Match 52.3%; Score 31.4; DB 1; Length 499;
 Best Local Similarity 77.6%; Pred. No. 12;
 Matches 38; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 8 CCGCACCGCCGACACCGGTCACAGCGCGACACTCGACCTGCG 56
 Db 295 CCGCACCGCCGACACCGGTCACAGAGACGACGATCACACG 247

RESULT 7

AUI94847 520 bp mRNA linear EST 14-OCT-2003
 LOCUS AUI94847 Porphyra yezoensis TU-1 sporophytes Porphyra yezoensis
 DEFINITION CDNA clone PFL066b09_r 5', mRNA sequence.
 ACCESSION AUI94847
 VERSION AUI94847.1 GI:31935887
 KEYWORDS EST.
 SOURCE Porphyra yezoensis
 ORGANISM Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae;

REFERENCE

1 (bases 1 to 520)
 Porphyra.

AUTHORS

Asamizu,E., Nakajima,M., Kitade,Y., Saga,N., Nakamura,Y. and
 Tabata,S.
 TITLE COMPARISON OF RNA EXPRESSION PROFILES BETWEEN THE TWO GENERATIONS
 OF PORPHYRA YEZOENSIS (RHODOPHYTA), BASED ON EXPRESSED SEQUENCE TAG
 FREQUENCY ANALYSIS
 J. Phycol. 39 (5), 923-930 (2003)
 CONTACT: Erika Asamizu
 The First Laboratory for Plant Gene Research
 Kazusa DNA Research Institute
 Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
 Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

JOURNAL

COMMENT The First Laboratory for Plant Gene Research
 Kazusa DNA Research Institute
 Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
 Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES

1..520
 location/Qualifiers
 /organism="Porphyra yezoensis"
 /mol_type="mRNA"
 /strain="TU-1"
 /db_xref="taxon:2788"
 /clone="PFL066b09_r"
 /dev_stage="sporophytes"
 /clone_lib="Porphyra yezoensis TU-1 sporophytes"

ORIGIN

Query Match 52.3%; Score 31.4; DB 1; Length 520;
 Best Local Similarity 77.6%; Pred. No. 12;
 Matches 38; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 8 CCGCACCGCCGACACCGGTCACAGCGCGACACTCGACCTGCG 56
 Db 209 CCGCACCGCCGACACCGGTCACAGAGACGACGATCACACG 161

RESULT 8

AUI94863 530 bp mRNA linear EST 14-OCT-2003
 LOCUS AUI94863 Porphyra yezoensis TU-1 sporophytes Porphyra yezoensis
 DEFINITION CDNA clone PFL066d09_r 5', mRNA sequence.
 ACCESSION AUI94863
 VERSION AUI94863.1 GI:31935919
 KEYWORDS EST.
 SOURCE Porphyra yezoensis
 ORGANISM Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae;

REFERENCE

1 (bases 1 to 530)
 Asamizu,E., Nakajima,M., Kitade,Y., Saga,N., Nakamura,Y. and
 Tabata,S.
 TITLE COMPARISON OF RNA EXPRESSION PROFILES BETWEEN THE TWO GENERATIONS
 OF PORPHYRA YEZOENSIS (RHODOPHYTA), BASED ON EXPRESSED SEQUENCE TAG
 FREQUENCY ANALYSIS
 J. Phycol. 39 (5), 923-930 (2003)
 CONTACT: Erika Asamizu
 The First Laboratory for Plant Gene Research
 Kazusa DNA Research Institute
 Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
 Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

JOURNAL

COMMENT The First Laboratory for Plant Gene Research
 Kazusa DNA Research Institute
 Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
 Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES

1..530
 location/Qualifiers
 /organism="Porphyra yezoensis"
 /mol_type="mRNA"
 /strain="TU-1"
 /db_xref="taxon:2788"
 /clone="PFL066d09_r"
 /dev_stage="sporophytes"
 /clone_lib="Porphyra yezoensis TU-1 sporophytes"

ORIGIN

Query Match 52.3%; Score 31.4; DB 1; Length 530;
 Best Local Similarity 77.6%; Pred. No. 12;
 Matches 38; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 8 CCGCACCGCCGACACCGGTCACAGCGCGACACTCGACCTGCG 56
 Db 209 CCGCACCGCCGACACCGGTCACAGAGACGACGATCACACG 161

RESULT 9
LOCUS AUI93189/c 552 bp mRNA linear EST 14-OCT-2003
DEFINITION AUI93189 Porphyra yezoensis TU-1 sporophytes Porphyra yezoensis
CDNA clone PFU041f12_r 5', mRNA sequence.
ACCESSION AUI93189
VERSION AUI93189.1 GI:31932583
KEYWORDS EST.
SOURCE Porphyra yezoensis
ORGANISM Porphyra yezoensis
Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae;
Porphyra.

REFERENCE 1 (bases 1 to 552)
AUTHORS Asamizu,E., Nakajima,M., Kitade,Y., Saga,N., Nakamura,Y. and
Tabata,S.
TITLE COMPARISON OF RNA EXPRESSION PROFILES BETWEEN THE TWO GENERATIONS
OF PORPHYRA YEZOENSIS (RHODOPHYTA), BASED ON EXPRESSED SEQUENCE TAG
FREQUENCY ANALYSIS
JOURNAL J. Phycol. 39 (5), 923-930 (2003)
COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.
Location/Qualifiers
1..552
/organism="Porphyra yezoensis"
/mol_type="mRNA"
/strain="TU-1"
/db_xref="taxon:2788"
/clone="PFU041f12_r"
/dev_stage="sporophytes"
/clone_1ib="Porphyra yezoensis TU-1 sporophytes"

ORIGIN
Query Match 52.3%; Score 31.4; DB 1; Length 552;
Best Local Similarity 77.6%; Pred. No. 12;
Matches 38; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

CY 8 CCGCACCGCCCGCACACGGGCTCACAGCGCCGACACATCGACCTGC 56
|||||
DB 302 CCGCACCGCCCGCACACGGGCTCACAGCGCCGACACATCGACCTGC 254
|||||

RESULT 10
LOCUS AI925867/c 330 bp mRNA linear EST 08-MAR-2000
DEFINITION wo20d04.x1 NCI_CGAP_Panl Homo sapiens cDNA clone IMAGE:2455879 3',
similar to contains element MSRI MSRI repetitive element ;, mRNA
sequence.
ACCESSION AI925867
VERSION AI925867
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 330)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Life Technologies catalog #: 11548-013
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert length: 578 Std Error: 0.00
Seq primer: -40UP from Gibco

JOURNAL COMMENT
TITLE
AUTHORS
REFERENCES
ORGANISM

High quality sequence stop: 305.
Location/Qualifiers
1..330
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2455879"
/tissue_type="adenocarcinoma"
/lab_host="DH10B"
/clone_1ib="NCI_CGAP_Panl"
/note="Organ: pancreas; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.72 kb. Life Technologies catalog #: 11548-013"

ORIGIN
Query Match 50.3%; Score 30.2; DB 1; Length 330;
Best Local Similarity 69.5%; Pred. No. 29;
Matches 41; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

CY 1 GGCACGACCGCCCGCCGACACGGGCTCACAGCGCCGACACATCGACTGGCC 59
|||||
DB 316 GGCCTCCACCGCCCGCCGACACGGGCTCACAGCGCCGACACATCGACTGGCC 258
|||||

RESULT 11
LOCUS BUI48487 1113 bp mRNA linear EST 03-SEP-2002
DEFINITION AGENCOURT 8670479 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6380642
5', mRNA sequence.
ACCESSION BUI48487
VERSION BUI48487.1 GI:22662019
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 1113)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: DCTP/DTP
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHCW2569 row: j column: 03
High quality sequence stop: 235.
Location/Qualifiers
1..1113
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6380642"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_1ib="NIH_MGC 40"
/note="Organ: prostate; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN
Query Match 50.3%; Score 30.2; DB 5; Length 1113;
Best Local Similarity 69.5%; Pred. No. 28;

Matches 41; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy 1 GGCAGACCGGACCGCCCGGACGCGGATCACAAAGCGCCGACAGACTCGACTGCGCC 59
Db 30 GGCTCCACCGCCCGCCCGACCGGATGTCACCTTGCCCGGACACCAAGCGCGCCCGCC 88

RESULT 12
BU542996
LOCUS
DEFINITION BU542996 1130 bp mRNA linear EST 13-SEP-2002
5', mRNA sequence.
ACCESSION BU542996
VERSION BU542996.1 GI:22853479
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NIH-MGC http://mgc.nci.nih.gov/
1 (bases 1 to 1130)
National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM2771 row: c column: 11
High quality sequence start: 27
High quality sequence stop: 246.
Location/Qualifiers
1. 1130
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6574979"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_1lb="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site: 1: XhoI;
Site 2: EcoRI; CDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC library."

ORIGIN

Query Match 50.3%; Score 30.2; DB 5; Length 1130;
Best local Similarity 69.5%; Pred. No. 28;
Matches 41; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy 1 GGCAGACCGGACCGCCCGGACGCGGATCACAAAGCGCCGACAGACTCGACTGCGCC 59
Db 69 GGCTCCACCGCCCGCCCGACCGGATGTCACCTTGCCCGGACACCAAGCGCGCCCGCC 127

RESULT 13
BO936898
LOCUS
DEFINITION BO936898 1234 bp mRNA linear EST 21-AUG-2002
5', mRNA sequence.
ACCESSION BO936898
VERSION BO936898.1 GI:22352281
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens (human)

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NIH-MGC http://mgc.nci.nih.gov/
1 (bases 1 to 1234)
National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM2577 row: c column: 24
High quality sequence stop: 245.
Location/Qualifiers
1. 1234
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6383855"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_1lb="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site: 1: XhoI;
Site 2: EcoRI; CDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC library."

ORIGIN

Query Match 50.3%; Score 30.2; DB 5; Length 1234;
Best local Similarity 69.5%; Pred. No. 28;
Matches 41; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy 1 GGCAGACCGGACCGCCCGGACGCGGATCACAAAGCGCCGACAGACTCGACTGCGCC 59
Db 30 GGCTCCACCGCCCGCCCGACCGGATGTCACCTTGCCCGGACACCAAGCGCGCCCGCC 88

RESULT 14
BO935496
LOCUS
DEFINITION BO935496 1262 bp mRNA linear EST 21-AUG-2002
5', mRNA sequence.
ACCESSION BO935496
VERSION BO935496.1 GI:22350879
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NIH-MGC http://mgc.nci.nih.gov/
1 (bases 1 to 1262)
National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM2573 row: f column: 09
High quality sequence start: 46
High quality sequence stop: 157.

This Page Blank (uspto)


```
XX SQ Sequence 60 BP; 13 A; 27 C; 17 G; 3 T; 0 U; 0 Other;
Query Match 100.0%; Score 60; DB 2; Length 60;
Best Local Similarity 100.0%; Pred. No. 7,3e-10;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCAGCAGCCGACCGCCGCGACACGCGGTCACAGCGCGCCAGACACTCGACCTGCGCCA 60
DB 1 GGCAGCAGCCGACCGCCGCGACACGCGGTCACAGCGCGCCAGACACTCGACCTGCGCCA 60

RESULT 2
AAV48322
ID AAV48322 standard; DNA; 60 BP.
XX
XX AAV48322;
AC
XX 20-NOV-1998 (first entry)
DT
XX
XX Nucleotide sequence encoding MUC1 tandem repeat unit R7.
DE
XX
XX ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
KM tumour; tumour-associated antigen.
XX
XX Homo sapiens.
OS
XX
XX WO9837095-A2.
PN
XX
XX 27-AUG-1998.
PD
XX
XX 24-FEB-1998; 98WO-US003693.
PF
XX
XX 24-FEB-1997; 97US-0038253P.
PR
XX
XX (THER-) THERION BIOLOGICS CORP.
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA (DAND ) DANA FARBEN CANCER INST INC.
XX
XX Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;
XX
XX WPI; 1998-467492/40.
DR
XX
XX New recombinant pox virus for tumour therapy - comprises DNA encoding an
PT immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
PT
XX
XX Disclosure; Page 11; 42pp; English.
PS
XX
XX The MUC1 tandem repeat units AAV48317-V48325 were used to create an
CC immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
CC (RPV). The RPV was used in a pharmaceutical composition also containing
CC an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The
CC recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
CC does not undergo significant genetic deletion, thereby providing an
CC unexpectedly stable and immunogenic pox virus. They can be used to
CC prevent or treat tumours expressing MUC1 tumour-associated antigens
XX
XX
XX Sequence 60 BP; 14 A; 24 C; 16 G; 6 T; 0 U; 0 Other;
QY Query Match 86.7%; Score 52; DB 2; Length 60;
Best Local Similarity 91.7%; Pred. No. 2.6e-07;
Matches 55; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
DB 1 GGCAGCAGCCGACCGCCGCGACACGCGGTCACAGCGCGCCAGACACTCGACCTGCGCCA 60
1 GGAAGTACCGCTCCACCTGCGACACGCGGTCACAGCGCGCCAGACACTCGACCTGCGCCA 60

RESULT 3
AAV48325
ID AAV48325 standard; DNA; 60 BP.
XX
XX AAV48325;
AC
```

```
XX DT 20-NOV-1998 (first entry)
XX
XX Nucleotide sequence encoding MUC1 tandem repeat unit R10.
DE
XX
XX ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
KM tumour; tumour-associated antigen.
XX
XX Homo sapiens.
OS
XX
XX WO9837095-A2.
PN
XX
XX 27-AUG-1998.
PD
XX
XX 24-FEB-1998; 98WO-US003693.
PF
XX
XX 24-FEB-1997; 97US-0038253P.
PR
XX
XX (THER-) THERION BIOLOGICS CORP.
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA (DAND ) DANA FARBEN CANCER INST INC.
XX
XX Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;
XX
XX WPI; 1998-467492/40.
DR
XX
XX New recombinant pox virus for tumour therapy - comprises DNA encoding an
PT immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
PT
XX
XX Disclosure; Page 11; 42pp; English.
PS
XX
XX The MUC1 tandem repeat units AAV48317-V48325 were used to create an
CC immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
CC (RPV). The RPV was used in a pharmaceutical composition also containing
CC an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The
CC recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
CC does not undergo significant genetic deletion, thereby providing an
CC unexpectedly stable and immunogenic pox virus. They can be used to
CC prevent or treat tumours expressing MUC1 tumour-associated antigens
XX
XX
XX Sequence 60 BP; 12 A; 22 C; 19 G; 7 T; 0 U; 0 Other;
QY Query Match 61.0%; Score 36.6; DB 2; Length 60;
Best Local Similarity 76.3%; Pred. No. 0.021;
Matches 45; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
DB 1 GGCAGCAGCCGACCGCCGCGACACGCGGTCACAGCGCGCCAGACACTCGACCTGCGGCC 59
1 GGTAGTACAGCGCGCACCGCCGCGACATGGCGTCCGAGTACGAGACCGGCGCC 59

RESULT 4
AAV48317
ID AAV48317 standard; DNA; 60 BP.
XX
XX AAV48317;
AC
XX
XX 20-NOV-1998 (first entry)
DT
XX
XX Nucleotide sequence encoding MUC1 tandem repeat unit R2.
DE
XX
XX ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
KM tumour; tumour-associated antigen.
XX
XX Homo sapiens.
OS
XX
XX WO9837095-A2.
PN
XX
XX 27-AUG-1998.
PD
XX
XX 24-FEB-1998; 98WO-US003693.
PF
XX
XX 24-FEB-1997; 97US-0038253P.
PR
```

```
XX (THER-) THERION BIOLOGICS CORP.
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA (DAND) DANA FARBBER CANCER INST INC.
XX
PI Schlom J, Kantor J, Kuife D, Panicali D, Gritz L;
XX WPI; 1998-467492/40.
DR
XX
PT New recombinant pox virus for tumour therapy - comprises DNA encoding an
PT immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
XX
PS Disclosure; Page 11; 42pp; English.
XX
CC The MUC1 tandem repeat units AAV48317-V48325 were used to create an
CC immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
CC (RPV). The RPV was used in a pharmaceutical composition also containing
CC an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The
CC recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
CC does not undergo significant genetic deletion, thereby providing an
CC unexpectedly stable and immunogenic pox virus. They can be used to
CC prevent or treat tumours expressing MUC1 tumour-associated antigens
XX
SQ Sequence 60 BP; 17 A; 21 C; 13 G; 9 T; 0 U; 0 Other;

Query Match          55.7%; Score 33.4; DB 2; Length 60;
Best Local Similarity 72.9%; Pred. No. 0.22;
Matches 43; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

OY      1 GGCAGACCGCGCCGCGGTGTCACAGGCGCCGACACTGCGTGGCC 59
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      1 GGCATCTCTGCACCGCACCGGACATGGCGTAACTCAGCAGCTGACAGACTGCGCC 59

RESULT 5
AAV48318
ID AAV48318 standard; DNA; 60 BP.
XX
AC AAV48318;
XX
DT 20-NOV-1998 (first entry)
XX
DE Nucleotide sequence encoding MUC1 tandem repeat unit R3.
XX
KW ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
KW tumour; tumour-associated antigen.
XX
OS Homo sapiens.
XX
PN WO9837095-A2.
XX
PD 27-AUG-1998.
XX
PS 24-FEB-1998; 98WO-US003693.
XX
PR 24-FEB-1997; 97US-0038253P.
XX
PA (THER-) THERION BIOLOGICS CORP.
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA (DAND) DANA FARBBER CANCER INST INC.
XX
PI Schlom J, Kantor J, Kuife D, Panicali D, Gritz L;
XX WPI; 1998-467492/40.
DR
XX
PT New recombinant pox virus for tumour therapy - comprises DNA encoding an
PT immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
XX
PS Disclosure; Page 11; 42pp; English.
XX
CC The MUC1 tandem repeat units AAV48317-V48325 were used to create an
CC immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
CC (RPV). The RPV was used in a pharmaceutical composition also containing
```

```
CC an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The
CC recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
CC does not undergo significant genetic deletion, thereby providing an
CC unexpectedly stable and immunogenic pox virus. They can be used to
CC prevent or treat tumours expressing MUC1 tumour-associated antigens
XX
SQ Sequence 60 BP; 7 A; 29 C; 19 G; 5 T; 0 U; 0 Other;

Query Match          53.0%; Score 31.8; DB 2; Length 60;
Best Local Similarity 71.2%; Pred. No. 0.7;
Matches 42; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

OY      1 GGCAGACCGCGCCGCGGTGTCACAGGCGCCGACACTGCGTGGCC 59
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      1 GCATCTCTGCACCGCACCGGACATGGCGTAACTCAGCAGCTGACAGACTGCGCC 59

RESULT 6
AAV48324
ID AAV48324 standard; DNA; 60 BP.
XX
AC AAV48324;
XX
DT 20-NOV-1998 (first entry)
XX
DE Nucleotide sequence encoding MUC1 tandem repeat unit R9.
XX
KW ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
KW tumour; tumour-associated antigen.
XX
OS Homo sapiens.
XX
PN WO9837095-A2.
XX
PD 27-AUG-1998.
XX
PS 24-FEB-1998; 98WO-US003693.
XX
PR 24-FEB-1997; 97US-0038253P.
XX
PA (THER-) THERION BIOLOGICS CORP.
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA (DAND) DANA FARBBER CANCER INST INC.
XX
PI Schlom J, Kantor J, Kuife D, Panicali D, Gritz L;
XX WPI; 1998-467492/40.
DR
XX
PT New recombinant pox virus for tumour therapy - comprises DNA encoding an
PT immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
XX
PS Disclosure; Page 11; 42pp; English.
XX
CC The MUC1 tandem repeat units AAV48317-V48325 were used to create an
CC immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
CC (RPV). The RPV was used in a pharmaceutical composition also containing
CC an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The
CC recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
CC does not undergo significant genetic deletion, thereby providing an
CC unexpectedly stable and immunogenic pox virus. They can be used to
CC prevent or treat tumours expressing MUC1 tumour-associated antigens
XX
SQ Sequence 60 BP; 12 A; 26 C; 14 G; 8 T; 0 U; 0 Other;

Query Match          51.0%; Score 30.6; DB 2; Length 60;
Best Local Similarity 73.6%; Pred. No. 1.7;
Matches 39; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

OY      7 ACCGACCGCGCGCGGTGTCACAGGCGCCGACACTGCGTGGCC 59
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      7 ACGGACCTCCGACGACGAGTCACTGTGACCGGACCGCGTCCGCTCC 59
```

```
RESULT 7
AAV48316
ID AAV48316 standard; cDNA; 60 BP.
XX
AC AAV48316;
XX
DT 20-NOV-1998 (first entry)
XX
DE Nucleotide sequence encoding MUC1 tandem repeat unit.
XX
KW ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
KW tumour; tumour-associated antigen.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT 1..60
FT CDS /*tag= a
FT /product= "MUC1 tandem repeat unit"
XX
PN WO9837095-A2.
XX
PD 27-AUG-1998.
XX
PF 24-FEB-1998; 98WO-US003693.
XX
PR 24-FEB-1997; 97US-0038253P.
XX
PA (THER-) THERION BIOLOGICS CORP.
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA (DAND) DANA FARBEN CANCER INST INC.
XX
PI Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;
XX
PI MPI; 1998-467492/40.
XX
DR P-PSDB; AAM77229.
XX
PT New recombinant pox virus for tumour therapy - comprises DNA encoding an
PT immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
XX
XX
XX Example 1; Page 20; 42pp; English.
XX
CC The MUC1 tandem repeat unit was used to create an immunogenic mini-MUC1
CC fragment for inclusion in a recombinant pox virus (RPV). The RPV was used
CC in a pharmaceutical composition also containing an immunomodulator to
CC generate MUC1 specific cytotoxic T-lymphocytes. The recombinant pox virus
CC therefore encodes an immunogenic MUC1 fragment that does not undergo
CC significant genetic deletion, thereby providing an unexpectedly stable
CC and immunogenic pox virus. They can be used to prevent or treat tumours
CC expressing MUC1 tumour-associated antigens
XX
SQ Sequence 60 BP; 7 A; 33 C; 16 G; 4 T; 0 U; 0 Other;
XX
Query Match 50.3%; Score 30.2; DB 2; Length 60;
Best Local Similarity 69.5%; Pred. No. 2.2;
Matches 41; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
XX
QY 1 GGCAGCAGCCGACCGCCGCGGATCAAGCGCGCCGACACTGACCTGCGCC 59
Db 1 GGCCTCAGCCGCCCCCGACCGCGGTGTCACTCGGCGCCCGGACACCGCGCGCCCC 59
XX
RESULT 8
AAV48326
ID AAV48326 standard; cDNA; 78 BP.
XX
AC AAV48326;
XX
DT 20-NOV-1998 (first entry)
XX
DE Nucleotide sequence encoding MUC1 tandem repeat unit a.
XX
KW ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
```

```
KW tumour; tumour-associated antigen.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT 1..78
FT CDS /*tag= a
FT /product= "MUC1 tandem repeat unit"
XX
PN WO9837095-A2.
XX
PD 27-AUG-1998.
XX
PF 24-FEB-1998; 98WO-US003693.
XX
PR 24-FEB-1997; 97US-0038253P.
XX
PA (THER-) THERION BIOLOGICS CORP.
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA (DAND) DANA FARBEN CANCER INST INC.
XX
PI Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;
XX
PI MPI; 1998-467492/40.
XX
DR P-PSDB; AAM77230.
XX
PT New recombinant pox virus for tumour therapy - comprises DNA encoding an
PT immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
XX
XX
XX Example 1; Page 20; 42pp; English.
XX
CC The MUC1 tandem repeat units AAV48326-V48328 were used to create an
CC immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
CC (RPV). The RPV was used in a pharmaceutical composition also containing
CC an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The
CC recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
CC does not undergo significant genetic deletion, thereby providing an
CC unexpectedly stable and immunogenic pox virus. They can be used to
CC prevent or treat tumours expressing MUC1 tumour-associated antigens
XX
SQ Sequence 78 BP; 9 A; 42 C; 22 G; 5 T; 0 U; 0 Other;
XX
Query Match 50.3%; Score 30.2; DB 2; Length 78;
Best Local Similarity 69.5%; Pred. No. 2.3;
Matches 41; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
XX
QY 1 GGCAGCAGCCGACCGCCGCGGATCAAGCGCGCCGACACTGACCTGCGCC 59
Db 1 GGCCTCAGCCGCCCCCGACCGCGGTGTCACTCGGCGCCCGGACACCGCGCGCCCC 59
XX
RESULT 9
AAN90579/C
ID AAN90579 standard; cDNA; 309 BP.
XX
AC AAN90579;
XX
DT 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 04-DEC-1989 (first entry)
XX
DE PDF9.3 cDNA insert.
XX
KW PDF9.3; human DF3 breast carcinoma-associated antigen epitope.
XX
OS Human MCF-7 breast carcinoma cells; 'ATCC HTB22'.
XX
PN WO8907107-A.
XX
PD 10-AUG-1989.
XX
PF 29-JAN-1988; 88US-00149831.
XX
```

PR 29-JAN-1988; 88US-00149831.
XX (DANA-) DANA-FARBER CANCER.
XX Kufe DW;
PI WPI; 1989-248989/34.
DR P-PSDB; AAP91045, AAP91053, AAP91054, AAP90146.
XX
PT Recombinant polypeptide(s) - contains DF 3 breast carcinoma antigen
PT epitope and useful as assay reagents, and encoding DNA sequences.
XX
PS Claim 1; Fig 4; 31pp; English.
CC The sequence encodes a carbohydrate-free polypeptide contg. DF3 breast
CC carcinoma antigen epitope. Useful as a competitive binding assay reagent
CC and improves diagnosis. The CDNA consists of nearly identical 60 BP
CC tandem repeats which are 85% GC-rich. See also AAP91045, AAP91046,
CC AAP91053, AAP91054. (Updated on 25-MAR-2003 to correct PA field.)
CC (Updated on 27-AUG-2003 to correct OS field.)
XX
SQ Sequence 309 BP; 21 A; 92 C; 160 G; 36 T; 0 U; 0 Other;
Query Match 50.3%; Score 30.2; DB 1; Length 309;
Best Local Similarity 69.5%; Pred. No. 2.4;
Matches 41; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
QY 1 GGCAGCAGCCGACCGCCGCGGATCACAAGCGCGCCAGACACTGAGCTGGCC 59
DB 266 GGCTCCAGCCGCCCCCGCCAGCCGATGTCACCTCGGCGCCGAGACAGCGCGGCC 208
RESULT 10
ID ADI57712 standard; cDNA; 1194 BP.
XX
AC ADI57712;
XX
DT 22-APR-2004 (first entry)
XX
DE Human breast specific nucleic acid (BSNA) #83.
XX
KW Human; breast specific nucleic acid; BSNA; gene; ss; metastasis;
KW breast cancer; cytostatic.
XX
OS Homo sapiens.
XX
PN WO2003106648-A2.
XX
PD 24-DEC-2003.
XX
PF 16-JUN-2003; 2003WO-US018934.
XX
PR 14-JUN-2002; 2002US-0389327P.
XX
PA (DIAD-) DIADEXUS INC.
XX
PI Salceda S, Macina RA, Turner LR, Sun Y, Liu C;
XX
DR WPI; 2004-082185/08.
DR P-PSDB; ADI57782.
XX
PT Novel isolated polypeptide comprising breast specific protein sequences,
PT useful for diagnosing or monitoring presence and metastases of breast
PT cancer in patient.
XX
PS Claim 1; SEQ ID NO 83; 370pp; English.
XX
CC The invention relates to human breast specific nucleic acids (BSNA) and
CC the breast specific proteins (BSP) they encode. The nucleic acids are
CC useful for determining the presence of a BSNA in a sample which involves
CC contacting the sample with a BSNA under conditions in which the BSNA will
CC selectively hybridise to a BSNA in the sample, and detecting the

CC hybridisation. The nucleic acids are useful for determining the presence
CC of a BSP in a sample which involves contacting the sample with suitable
CC reagent under conditions in which the reagent will selectively interact
CC with the BSP, and detecting the interaction of the reagent with a BSP in
CC the sample. The nucleic acids and proteins are useful for diagnosing or
CC monitoring the presence and metastases of breast cancer in a patient,
CC which involves determining an amount of nucleic acid or protein in and
CC comparing the determined amount of nucleic acid or protein in the sample
CC of the patient to the amount of a breast specific marker in a normal
CC control, where a difference in the determined amount in the sample
CC compared to the amount in the control is associated with the presence of
CC breast cancer. The sequences are useful for treating a patient with
CC breast cancer, involving administering a composition consisting of a BSNA
CC or a BSP to a patient, where the administration induces an immune
CC response against the breast cancer cell expressing the BSNA or BSP. This
CC sequence represents a human BSNA of the invention.
XX
SQ Sequence 1194 BP; 243 A; 438 C; 275 G; 238 T; 0 U; 0 Other;
Query Match 50.3%; Score 30.2; DB 12; Length 1194;
Best Local Similarity 69.5%; Pred. No. 2.6;
Matches 41; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
QY 1 GGCAGCAGCCGACCGCCGCGGATCACAAGCGCGCCAGACACTGAGCTGGCC 59
DB 596 GGCTCCAGCCGCCCCCGCCAGCCGATGTCACCTCGGCGCCGAGACAGCGCGGCC 654
RESULT 11
ID ADI57693 standard; cDNA; 1378 BP.
XX
AC ADI57693;
XX
DT 22-APR-2004 (first entry)
XX
DE Human breast specific nucleic acid (BSNA) #64.
XX
KW Human; breast specific nucleic acid; BSNA; gene; ss; metastasis;
KW breast cancer; cytostatic.
XX
OS Homo sapiens.
XX
PN WO2003106648-A2.
XX
PD 24-DEC-2003.
XX
PF 16-JUN-2003; 2003WO-US018934.
XX
PR 14-JUN-2002; 2002US-0389327P.
XX
PA (DIAD-) DIADEXUS INC.
XX
PI Salceda S, Macina RA, Turner LR, Sun Y, Liu C;
XX
DR WPI; 2004-082185/08.
DR P-PSDB; ADI57765.
XX
PT Novel isolated polypeptide comprising breast specific protein sequences,
PT useful for diagnosing or monitoring presence and metastases of breast
PT cancer in patient.
XX
PS Claim 1; SEQ ID NO 64; 370pp; English.
XX
CC The invention relates to human breast specific nucleic acids (BSNA) and
CC the breast specific proteins (BSP) they encode. The nucleic acids are
CC useful for determining the presence of a BSNA in a sample which involves
CC contacting the sample with a BSNA under conditions in which the BSNA will
CC selectively hybridise to a BSNA in the sample, and detecting the
CC hybridisation. The nucleic acids are useful for determining the presence
CC of a BSP in a sample which involves contacting the sample with suitable
CC reagent under conditions in which the reagent will selectively interact
CC with the BSP, and detecting the interaction of the reagent with a BSP in

CC the sample. The nucleic acids and proteins are useful for diagnosing or
CC monitoring the presence and metastases of breast cancer in a patient,
CC which involves determining an amount of nucleic acid or protein and
CC comparing the determined amount of nucleic acid or protein in the sample
CC of the patient to the amount of a breast specific marker in a normal
CC control, where a difference in the determined amount in the sample
CC compared to the amount in the control is associated with the presence of
CC breast cancer. The sequences are useful for treating a patient with
CC or a BSP to a patient, where the administration induces an immune
CC response against the breast cancer cell expressing the BSNA or BSP. This
CC sequence represents a human BSNA of the invention.

CC Sequence 1378 BP; 295 A; 489 C; 316 G; 278 T; 0 U; 0 Other;

Query Match 50.3%; Score 30.2; DB 12; Length 1378;
Best Local Similarity 69.5%; Pred. No. 2.6;
Matches 41; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

OY 1 GGACAGACCGCGCGCGGAGTCAAGCGCGCGACAGACTGACCTGCGGCC 59
Db 596 GGCTCCACCGCGCGCGCGCGGAGTCAAGCGCGCGCGACAGCGCGCGGCC 654

RESULT 12
ADO23180/c
ID ADO23180 standard; RNA; 1424 BP.

AC ADO23180;
DT 12-AUG-2004 (first entry)

DE Antisense human MUC1 mucin glycoprotein RNA (coding sequence) SeqID 75.

KW human; MUC1; mucin glycoprotein; cancer; chemotherapeutic; MUC1/ECG; ss;
KW antisense.

OS Homo sapiens.

XX WO2004044160-A2.

XX 27-MAY-2004.

PD 12-NOV-2003; 2003WO-US035848.

PR 13-NOV-2002; 2002US-00293391.

PR 29-MAY-2003; 2003US-00447839.

PA (DAND) DANA FARBER CANCER INST INC.
PA (ILEX-) ILEX PROD INC.

PI Kufe DW, Kharbanda S, Weisman SD;
PI WPI; 2004-420304/39.

XX DR

PT Double-stranded RNA complex useful for inhibiting proliferation of cancer
PT cell expressing MUC1 mucin glycoprotein, comprises first and second RNA
PT sequences.

XX PS

PS Disclosure; SEQ ID NO 75; 112bp; English.

CC This invention relates to novel modulators of the human MUC1 mucin
CC glycoprotein for use in cancer therapeutics, where MUC1 is a protein that
CC acts to inhibit the apoptotic response to genotoxic stress caused by
CC chemotherapeutic agents. In particular, it refers to modulators of the
CC MUC1 extracellular domain (MUC1/ECG). The method refers to using double-
CC stranded RNA complexes as MUC1 interference RNA compositions such that
CC MUC1 expression is inhibited, which in turn inhibits cancer cell
CC proliferation. The present invention describes screening assays to
CC identify compounds that inhibit the binding of various MUC1 ligands such
CC as neutrogin 2, as well as agonists, antagonists and antibodies thereof.
CC Furthermore, it provides MUC1 antisense, small interfering RNA (siRNA)
CC and small molecules in combination with chemotherapeutic agents that are

CC useful in the field of cancer therapy. This polynucleotide sequence is
CC the antisense human MUC1 RNA of the invention.

CC Sequence 1424 BP; 309 A; 325 C; 495 G; 0 T; 295 U; 0 Other;

Query Match 50.3%; Score 30.2; DB 12; Length 1424;
Best Local Similarity 69.5%; Pred. No. 2.6;
Matches 41; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

OY 1 GGACAGACCGCGCGCGGAGTCAAGCGCGCGCGACAGACTGACCTGCGGCC 59
Db 1041 GGCTCCACCGCGCGCGCGCGGAGTCAAGCGCGCGCGACAGCGCGCGGCC 983

RESULT 13
ABL60159
ID ABL60159 standard; cDNA; 1428 BP.

AC ABL60159;

DT 22-JUL-2002 (first entry)

DE Human MUC1 encoding cDNA SEQ ID NO 2.

KW Human; mucin 1; MUC1; transmembrane protein; SNP; cancer; cytostatic;
KW single nucleotide polymorphism; haplotyping; genotyping; drug;
KW antiinflammatory; gene; ss.

XX OS Homo sapiens.

XX Key

FT location/Qualifiers
FT 1..1428
FT /*tag= a
FT /product= "MUC1"
FT replace(1009,A)
FT /*tag= b
FT /standard name= "Single nucleotide polymorphism"
FT /note= "SNP allelic variation results in Val substituted
FT by Met at position 337 of the MUC1 protein (ABB77476)"

XX variation

XX WO200226765-A2.

XX 04-APR-2002.

PD 25-SEP-2001; 2001WO-US030151.

PR 28-SEP-2000; 2000US-0236113P.

PR (GENA-) GENAISSANCE PHARM INC.

PA Chew A, Koshy B;
PA WPI; 2002-405042/43.

PI P-PSDB; ABB77476.

XX DR

PT New genetic variants of mucin 1, Transmembrane gene, useful in studying
PT expression and function of protein encoded by the gene and for screening
PT drugs to treat diseases e.g. cancer.

XX PS

PS Claim 23; Fig 2; 75pp; English.

CC The invention relates to a polynucleotide (ABL60158, ABL60159) encoding
CC mucin 1/MUC1 (ABB77476), Transmembrane isogene. The invention describes
CC novel genetic variants of the MUC1 gene. The invention is useful for
CC haplotyping/genotyping the MUC1 gene in an individual and identifying an
CC association between a trait and at least one of the haplotypes or
CC haplotype pairs of MUC1 gene. MUC1 is useful for studying the expression
CC and function of MUC1 and expressing MUC1 protein for use in screening for
CC candidate drugs to treat diseases related to MUC1 activity and in
CC studying the effect of the variation on the biological activity of MUC1
CC as well as on the binding affinity of candidate drugs targeting MUC1 for
CC the treatment of e.g. cancer. MUC1 is further used by the pharmaceutical
CC research scientist to validate MUC1 as a candidate target for and in

CC design of clinical trials of candidate drugs for, treating a specific
 CC condition drugs or disease predicted to be associated with MUC1 activity.
 CC MUC1 antibodies are useful in a variety of diagnostic and prognostic
 CC formats and therapeutic methods

XX Sequence 1428 BP; 296 A; 496 C; 327 G; 309 T; 0 U; 0 Other;

Query Match 50.3%; Score 30.2; DB 6; Length 1428;
 Best Local Similarity 69.5%; Pred. No. 2.6;
 Matches 41; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1 GGCAGACCGCACCGCCCGCACAGGGGTCAAGAGCGCGACAGACTGACCTGGCC 59
 DB 385 GGCTCACCAGCCCCCCCCACAGCCACGCTGTACCTCGGCCCCGAGACACAGCGCGCCCC 443

RESULT 14

ADO23125
 ID ADO23125 standard; RNA; 1428 BP.

XX ADO23125;

DT 12-AUG-2004 (first entry)

DE Human MUC1 mucin glycoprotein RNA (coding sequence) SegID 20.

XX human; MUC1; mucin glycoprotein; cancer; chemotherapeutic; MUC1/BCD; ss.

OS Homo sapiens.

PN WO200404160-A2.

XX 27-MAY-2004.

PF 12-NOV-2003; 2003WO-US035848.

XX 13-NOV-2002; 2002US-00293391.

PR 29-MAY-2003; 2003US-00447839.

XX (DAND) DANA FARBER CANCER INST INC.

PA (ILEX-) ILEX PROD INC.

PI Kufe DW, Kharbanda S, Weltman SD;

DR WPI; 2004-420304/39.

PT Double-stranded RNA complex useful for inhibiting proliferation of cancer
 cell expressing MUC1 mucin glycoprotein, comprises first and second RNA
 sequences.

XX Claim 2; SEQ ID NO 20; 112pp; English.

XX This invention relates to novel modulators of the human MUC1 mucin
 glycoprotein for use in cancer therapeutics, where MUC1 is a protein that
 acts to inhibit the apoptotic response to genotoxic stress caused by
 chemotherapeutic agents. In particular, it refers to modulators of the
 MUC1 extracellular domain (MUC1/BCD). The method refers to using double-
 stranded RNA complexes as MUC1 interference RNA compositions such that
 MUC1 expression is inhibited, which in turn inhibits cancer cell
 proliferation. The present invention describes screening assays to
 identify compounds that inhibit the binding of various MUC1 ligands such
 as neutrophil 2, as well as agonists, antagonists and antibodies thereof.
 Furthermore, it provides MUC1 antisense, small interfering RNA (siRNA)
 and small molecules in combination with chemotherapeutic agents that are
 useful in the field of cancer therapy. This polynucleotide sequence is
 the human MUC1 RNA of the invention.

XX Sequence 1428 BP; 296 A; 496 C; 327 G; 0 T; 309 U; 0 Other;

Query Match 50.3%; Score 30.2; DB 12; Length 1428;
 Best Local Similarity 67.8%; Pred. No. 2.6;
 Matches 40; Conservative 1; Mismatches 18; Indels 0; Gaps 0;

QY 1 GGCAGACCGCACCGCCCGCACAGGGGTCAAGAGCGCGACAGACTGACCTGGCC 59
 DB 385 GGCTCACCAGCCCCCCCCACAGCCACGCTGTACCTCGGCCCCGAGACACAGCGCGCCCC 443

RESULT 15

ADF32627
 ID ADF32627 standard; DNA; 1457 BP.

XX ADF32627;

DT 26-FEB-2004 (first entry)

DE Plasmid JMW358 MUC-1 nucleotide sequence.

XX MUC-1 antigen; immune response; MUC-1; variable number of tandem repeat;
 KW VNTN; repeat unit; tumour; metastasis; cytostatic; vaccine; gene therapy;
 gene; ds.

XX Synthetic.

OS Synthetic.

PN WO200310060-A2.

XX 04-DEC-2003.

PF 23-MAY-2003; 2003WO-EP005594.

PR 24-MAY-2002; 2002GB-00012046.

XX (GLAXO) GLAXO GROUP LTD.

PA Burden N, Ellis JH, Hamblin PA;

PI WPI; 2004-042811/04.

PT New nucleic acid molecule encoding a MUC-1 antigen, useful for preparing
 a composition for treating or preventing tumors or metastases.

XX Example; Fig 3; 6pp; English.

XX The present invention describes a nucleic acid molecule which encodes a
 MUC-1 antigen. The nucleic acid is capable of raising an immune response
 in vivo, has reduced susceptibility to recombination than full-length MUC
 -1 and comprises between 1 and 15 variable number of tandem repeats
 (VNTN) perfect repeat units. Also described: (1) a plasmid comprising the
 CC DNA molecule; (2) a protein encoded by the nucleic acid; (3) a
 CC pharmaceutical composition comprising the nucleic acid; plasmid or
 CC protein and an excipient, diluent or carrier; and (4) a method of
 CC treating or preventing tumors or metastases. A MUC1 antigen has
 CC cytostatic activity, and can be used in vaccines, and in gene therapy.
 CC The nucleic acid is useful for preparing a composition for treating or
 CC preventing tumors or metastases. The present sequence is used in the
 CC exemplification of the present invention.

XX Sequence 1457 BP; 302 A; 506 C; 334 G; 315 T; 0 U; 0 Other;

Query Match 50.3%; Score 30.2; DB 12; Length 1457;
 Best Local Similarity 69.5%; Pred. No. 2.6;
 Matches 41; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1 GGCAGACCGCACCGCCCGCACAGGGGTCAAGAGCGCGACAGACTGACCTGGCC 59
 DB 403 GGCTCACCAGCCCCCCCCACAGCCACGCTGTACCTCGGCCCCGAGACACAGCGCGCCCC 461

Search completed: January 15, 2005, 20:36:10
 Job time : 171.5 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: January 15, 2005, 22:37:00 ; Search time 183.4 Seconds
(without alignments)
1879.789 Million cell updates/sec

Title: US-10-057-136-7

Perfect score: 60
Sequence: 1 ggttcagcggccccctcgc.....cggataccagcggccccct 60

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 4300275 seqs, 2872944193 residues

Total number of hits satisfying chosen parameters: 8600550

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Published Applications NA.*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
21: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	60	100.0	60	US-10-057-136-7	Sequence 7, Appl1
2	60	100.0	2297	US-10-406-317-41	Sequence 41, Appl1
3	43	71.7	60	US-10-057-136-2	Sequence 2, Appl1
4	43	71.7	1424	US-10-447-839A-75	Sequence 75, Appl1
5	43	71.7	1428	US-10-447-839A-70	Sequence 20, Appl1
6	43	71.7	1527	US-10-057-136-19	Sequence 19, Appl1
7	43	71.7	1793	US-10-447-839A-19	Sequence 19, Appl1
8	43	71.7	1804	US-09-864-824A-573	Sequence 573, Appl1
9	43	71.7	1804	US-10-029-517-17	Sequence 17, Appl1
10	43	71.7	1804	US-10-717-597-30	Sequence 30, Appl1
11	43	71.7	1804	US-10-775-920-84	Sequence 84, Appl1
12	43	71.7	1823	US-10-101-510-339	Sequence 339, Appl1

13	43	71.7	4139	US-09-964-824A-105	Sequence 105, App
14	43	71.7	4139	US-09-964-824A-578	Sequence 578, App
15	43	71.7	4139	US-09-864-864-334	Sequence 334, App
16	43	71.7	4139	US-09-880-107-2121	Sequence 2121, App
17	43	71.7	4139	US-09-968-007A-751	Sequence 751, App
18	43	71.7	4139	US-10-171-311-157	Sequence 157, App
19	43	71.7	4139	US-10-177-293-310	Sequence 310, App
20	43	71.7	4139	US-10-440-464-155	Sequence 155, App
21	43	71.7	4139	US-10-734-564-53	Sequence 53, App1
22	43	71.7	4139	US-10-775-920-80	Sequence 80, App1
23	43	71.7	4139	US-10-775-920-85	Sequence 85, App1
24	43	71.7	8181	US-10-447-839A-18	Sequence 18, App1
25	43	71.7	8186	US-10-247-703-23	Sequence 23, App1
26	43	71.7	8186	US-10-029-517-19	Sequence 19, App1
27	41.4	69.0	572	US-10-029-517-18	Sequence 18, App1
28	41.4	69.0	572	US-10-775-920-93	Sequence 93, App1
29	41.4	69.0	1721	US-09-864-864-280	Sequence 280, App
30	41.4	69.0	1721	US-09-967-768A-224	Sequence 224, App
31	41.4	69.0	1721	US-10-247-703-21	Sequence 21, App1
32	41.4	69.0	1721	US-10-097-340-211	Sequence 211, App
33	41.4	69.0	1721	US-10-171-311-155	Sequence 155, App
34	41.4	69.0	1721	US-10-007-926A-58	Sequence 58, App1
35	41.4	69.0	1721	US-10-029-517-3	Sequence 3, App1
36	41.4	69.0	1721	US-10-172-118-775	Sequence 775, App
37	41.4	69.0	1721	US-10-342-887-775	Sequence 775, App
38	41.4	69.0	1721	US-10-775-920-88	Sequence 88, App1
39	41.4	69.0	2026	US-10-198-846-12589	Sequence 12589, A
40	41.4	69.0	2238	US-10-775-920-87	Sequence 87, App1
41	41.4	69.0	2678	US-10-252-157-103	Sequence 103, App
42	41	68.3	518	US-10-247-703-38	Sequence 38, App1
43	41	68.3	518	US-10-029-517-101	Sequence 101, App
44	39.8	66.3	78	US-10-057-136-13	Sequence 13, App1
45	39.8	66.3	3343	US-10-247-703-27	Sequence 27, App1

ALIGNMENTS

RESULT 1
US-10-057-136-7
; Sequence 7, Application US/10057136
; Publication No. US20030021770A1
GENERAL INFORMATION:
; APPLICANT: SCHLON, JEFFREY
; APPLICANT: KANTOR, JUDITH
; APPLICANT: KUFFE, DONALD
; APPLICANT: PANICALI, DENNIS
; APPLICANT: GRITZ, LINDA
; TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
; TITLE OF INVENTION: TUMOR-ASSOCIATED ANTIGEN
; FILE REFERENCE: 700953/47113C
; CURRENT APPLICATION NUMBER: US/10/057,136
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/366,670
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: PCT/US98/03693
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/038,253
; PRIOR FILING DATE: 1997-02-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-136-7

Query Match 100.0%; Score 60; DB 14; Length 60;
Best Local Similarity 100.0%; Pred. No. 1,4e12;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 ggttcagcggccccctcgcagcgtgaaacatccgccccgatatcagacgccccct 60
|||||

Db 1 GGTTGACGCGCCCCCTGCTCAAGGTGTACATCCGCCCGGATACAGACCGGCCCT 60

RESULT 2

US-10-406-317-41
; Sequence 41, Application US/10406317
; Publication No. US20040019195A1
; GENERAL INFORMATION:
; APPLICANT: Schlom, Jeffrey;
; APPLICANT: Hodge, James;
; APPLICANT: Panticall, Dennis
; TITLE OF INVENTION: A recombinant vector expressing multiple constitutively
; FILE REFERENCE: 38163-0189
; CURRENT APPLICATION NUMBER: US/10/406.317
; CURRENT FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: US/09/856.988
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: PCT/US99/26866
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/111,582
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41
; LENGTH: 2297
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: VECTOR
US-10-406-317-41

Query Match 100.0%; Score 60; DB 16; Length 2297;
Best Local Similarity 100.0%; Pred. No. 1.2e-12;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTGACGCGCCCCCTGCTCAAGGTGTACATCCGCCCGGATACAGACCGGCCCT 60
Db 586 GGTTGACGCGCCCCCTGCTCAAGGTGTACATCCGCCCGGATACAGACCGGCCCT 645

RESULT 3

US-10-057-136-2
; Sequence 2, Application US/10057136
; Publication No. US20030021770A1
; GENERAL INFORMATION:
; APPLICANT: SCHLOM, JEFFREY
; APPLICANT: KANTOR, JUDITH
; APPLICANT: KUFE, DONALD
; APPLICANT: PANICALL, DENNIS
; APPLICANT: GRITZ, LINDA
; TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
; FILE REFERENCE: 700953/47113C
; CURRENT APPLICATION NUMBER: US/10/057.136
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/366.670
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: PCT/US98/03693
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/038,253
; PRIOR FILING DATE: 1997-02-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-136-2

Query Match 71.7%; Score 43; DB 14; Length 60;
Best Local Similarity 83.1%; Pred. No. 2.6e-06;

Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 GGTTGACGCGCCCCCTGCTCAAGGTGTACATCCGCCCGGATACAGACCGGCCCT 59
Db 1 GGTTGACGCGCCCCCTGCTCAAGGTGTACATCCGCCCGGATACAGACCGGCCCT 59

RESULT 4

US-10-447-839A-75/c
; Sequence 75, Application US/10447839A
; Publication No. US20040018181A1
; GENERAL INFORMATION:
; APPLICANT: Kufe, Donald W.
; APPLICANT: Kharbada, Suresder
; APPLICANT: Weltman, Steven D.
; TITLE OF INVENTION: MUC1 INTERFERENCE RNA COMPOSITIONS AND METHODS DERIVED THEREFROM
; FILE REFERENCE: 1000.05.009
; CURRENT APPLICATION NUMBER: US/10/447.839A
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: 10/293,391
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 09/951,938
; PRIOR FILING DATE: 2001-09-11
; PRIOR APPLICATION NUMBER: 60/231,841
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 75
; LENGTH: 1424
; TYPE: RNA
; ORGANISM: ARTIFICIAL
; FEATURE:
; OTHER INFORMATION: Synthesized Sequence
US-10-447-839A-75

Query Match 71.7%; Score 43; DB 16; Length 1424;
Best Local Similarity 83.1%; Pred. No. 2.3e-06;
Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 GGTTGACGCGCCCCCTGCTCAAGGTGTACATCCGCCCGGATACAGACCGGCCCT 59
Db 1041 GGTTGACGCGCCCCCTGCTCAAGGTGTACATCCGCCCGGATACAGACCGGCCCT 983

RESULT 5

US-10-447-839A-20
; Sequence 20, Application US/10447839A
; Publication No. US20040018181A1
; GENERAL INFORMATION:
; APPLICANT: Kufe, Donald W.
; APPLICANT: Kharbada, Suresder
; APPLICANT: Weltman, Steven D.
; TITLE OF INVENTION: MUC1 INTERFERENCE RNA COMPOSITIONS AND METHODS DERIVED THEREFROM
; FILE REFERENCE: 1000.05.009
; CURRENT APPLICATION NUMBER: US/10/447.839A
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: 10/293,391
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 09/951,938
; PRIOR FILING DATE: 2001-09-11
; PRIOR APPLICATION NUMBER: 60/231,841
; PRIOR FILING DATE: 2000-09-11
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 20
; LENGTH: 1428
; TYPE: RNA
; ORGANISM: RNA
US-10-447-839A-20

Query Match 71.7%; Score 43; DB 16; Length 1428;
Best Local Similarity 76.3%; Pred. No. 2.3e-06;
Matches 45; Conservative 4; Mismatches 10; Indels 0; Gaps 0;


```
/ Publication No. US20040110221A1
/ GENERAL INFORMATION:
/ APPLICANT: Wyeth
/ APPLICANT: Butczynski, Michael E.
/ APPLICANT: Twine, Natalie C.
/ APPLICANT: Dornier, Andrew J.
/ APPLICANT: Trepicchio, William L.
/ APPLICANT: Slonim, Donna K.
/ APPLICANT: Stover, Jennifer A.
/ TITLE OF INVENTION: METHODS FOR DIAGNOSING RCC AND OTHER SOLID TUMORS
/ FILE REFERENCE: AM101080L
/ CURRENT APPLICATION NUMBER: US/10/717,597
/ CURRENT FILING DATE: 2003-11-21
/ PRIOR APPLICATION NUMBER: US 60/459,782
/ PRIOR FILING DATE: 2003-04-03
/ PRIOR APPLICATION NUMBER: US 60/427,982
/ PRIOR FILING DATE: 2002-11-21
/ NUMBER OF SEQ ID NOS: 4904
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 30
/ LENGTH: 1804
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-717-597-30
```

```
Query Match      71.7%; Score 43; DB 17; Length 1804;
Best Local Similarity 83.1%; Pred. No. 2.2e-06;
Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
```

```
QY 1 GGTTGACGCGCCCCCTGCTACGGTGTAACTCCGCCCGGATACGACCGCGCCCC 59
Db 457 GGCTCCACCGCCCCCTGCTACGGTGTAACTCCGCCCGGATACGACCGCGCCCC 515
```

```
RESULT 11
US-10-775-920-84
/ Sequence 84, Application US/10775920
/ Publication No. US2004017544A1
/ GENERAL INFORMATION:
/ APPLICANT: Merzen Ltd
/ TITLE OF INVENTION: POLYNUCLEOTIDE SEQUENCES AND CORRESPONDING ENCODED POLYPEPTIDES
/ TITLE OF INVENTION: OF PARTICULAR SECRETED AND MEMBRANE-BOUND PROTEINS OVEREXPRESSED
/ FILE REFERENCE: Merzen - 0010B
/ CURRENT APPLICATION NUMBER: US/10/775,920
/ CURRENT FILING DATE: 2004-02-10
/ PRIOR APPLICATION NUMBER: US 60/447,900
/ PRIOR FILING DATE: 2003-02-13
/ NUMBER OF SEQ ID NOS: 385
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 84
/ LENGTH: 1804
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-775-920-84
```

```
Query Match      71.7%; Score 43; DB 17; Length 1804;
Best Local Similarity 83.1%; Pred. No. 2.2e-06;
Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
```

```
QY 1 GGTTGACGCGCCCCCTGCTACGGTGTAACTCCGCCCGGATACGACCGCGCCCC 59
Db 457 GGCTCCACCGCCCCCTGCTACGGTGTAACTCCGCCCGGATACGACCGCGCCCC 515
```

```
RESULT 12
US-10-101-510-339
/ Sequence 339, Application US/10101510
/ Publication No. US20030148295A1
/ GENERAL INFORMATION:
/ APPLICANT: WAN, JACKSON
/ APPLICANT: WANG, YIXIN
/ TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
```

```
/ FILE REFERENCE: 15117.0012
/ CURRENT APPLICATION NUMBER: US/10/101,510
/ CURRENT FILING DATE: 2002-03-20
/ PRIOR APPLICATION NUMBER: 60/276,947
/ PRIOR FILING DATE: 2001-03-20
/ NUMBER OF SEQ ID NOS: 805
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 339
/ LENGTH: 1823
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-101-510-339
```

```
Query Match      71.7%; Score 43; DB 15; Length 1823;
Best Local Similarity 83.1%; Pred. No. 2.2e-06;
Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
```

```
QY 1 GGTTGACGCGCCCCCTGCTACGGTGTAACTCCGCCCGGATACGACCGCGCCCC 59
Db 457 GGCTCCACCGCCCCCTGCTACGGTGTAACTCCGCCCGGATACGACCGCGCCCC 515
```

```
RESULT 13
US-09-964-824A-105
/ Sequence 105, Application US/09964824A
/ Patent No. US20020102531A1
/ GENERAL INFORMATION:
/ APPLICANT: Horrigan, Stephen
/ TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
/ TITLE OF INVENTION: Sets
/ FILE REFERENCE: 689290-73
/ CURRENT APPLICATION NUMBER: US/09/964,824A
/ CURRENT FILING DATE: 2001-09-27
/ PRIOR APPLICATION NUMBER: US/60/236,033
/ PRIOR FILING DATE: 2000-09-28
/ PRIOR APPLICATION NUMBER: US/60/236,032
/ PRIOR FILING DATE: 2000-09-28
/ PRIOR APPLICATION NUMBER: US/60/236,028
/ NUMBER OF SEQ ID NOS: 583
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 105
/ LENGTH: 4139
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-964-824A-105
```

```
Query Match      71.7%; Score 43; DB 9; Length 4139;
Best Local Similarity 83.1%; Pred. No. 2.2e-06;
Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
```

```
QY 1 GGTTGACGCGCCCCCTGCTACGGTGTAACTCCGCCCGGATACGACCGCGCCCC 59
Db 458 GGCTCCACCGCCCCCTGCTACGGTGTAACTCCGCCCGGATACGACCGCGCCCC 516
```

```
RESULT 14
US-09-964-824A-578
/ Sequence 578, Application US/09964824A
/ Patent No. US20020102531A1
/ GENERAL INFORMATION:
/ APPLICANT: Horrigan, Stephen
/ TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
/ FILE REFERENCE: 689290-73
/ CURRENT APPLICATION NUMBER: US/09/964,824A
/ CURRENT FILING DATE: 2001-09-27
/ PRIOR APPLICATION NUMBER: US/60/236,033
/ PRIOR FILING DATE: 2000-09-28
/ PRIOR APPLICATION NUMBER: US/60/236,032
/ PRIOR FILING DATE: 2000-09-28
/ PRIOR APPLICATION NUMBER: US/60/236,028
/ PRIOR FILING DATE: 2000-09-28
```

NUMBER OF SEQ ID NOS: 583
SOFTWARE: Patentin version 3.0
SEQ ID NO 578
LENGTH: 4139
TYPE: DNA
ORGANISM: Homo sapiens
US-09-964-824A-578

Query Match 71.7%; Score 43; DB 9; Length 4139;
Best Local Similarity 83.1%; Pred. No. 2.2e-06;
Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 GGTTCGACGCGCCCTCTCTACGCGTGTAAATCCGCGGATACAGACCGGCCCC 59
|||
Db 458 GGCTCCACCGCCCGCCCGACGCGTGTACCTCGCGCCCGACACAGCGCGCCCC 516

RESULT 15

US-09-864-864-334
Sequence 334, Application US/09864864
Patent No. US20020102679A1
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan L.
APPLICANT: Dillon, David C.
APPLICANT: Secrist, Heather
APPLICANT: Lodes, Michael J.
APPLICANT: Algate, Paul A.
APPLICANT: Fling, Steve P.
APPLICANT: Mannion, Jane
APPLICANT: Benson, Darin R.
APPLICANT: Carter, Darick
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
FILE REFERENCE: 210121.523
CURRENT APPLICATION NUMBER: US/09/864,864
CURRENT FILING DATE: 2001-05-23
NUMBER OF SEQ ID NOS: 341
SOFTWARE: Corixa Invention Disclosure Database
SEQ ID NO 334
LENGTH: 4139
TYPE: DNA
ORGANISM: Homo sapiens
US-09-864-864-334

Query Match 71.7%; Score 43; DB 9; Length 4139;
Best Local Similarity 83.1%; Pred. No. 2.2e-06;
Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 GGTTCGACGCGCCCTCTCTACGCGTGTAAATCCGCGGATACAGACCGGCCCC 59
|||
Db 458 GGCTCCACCGCCCGCCCGACGCGTGTACCTCGCGCCCGACACAGCGCGCCCC 516

Search completed: January 16, 2005, 09:30:25
Job time : 184.4 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2005, 19:50:13 ; Search time 35.2 Seconds
(without alignments)
1211.572 Million cell updates/sec

Title: US-10-057-136-7

Perfect score: 60

Sequence: 1 ggttcagcggcccccctgc.....cgataccagaccgcccct 60

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/prodata/1/ina/5A.COMB.seq: *
2: /cgn2_6/prodata/1/ina/5B.COMB.seq: *
3: /cgn2_6/prodata/1/ina/6A.COMB.seq: *
4: /cgn2_6/prodata/1/ina/6B.COMB.seq: *
5: /cgn2_6/prodata/1/ina/6CTUS.COMB.seq: *
6: /cgn2_6/prodata/1/ina/Backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	43	71.7	1804	US-10-029-517-17	Sequence 17, Appl
2	43	71.7	8186	US-10-029-517-19	Sequence 19, Appl
3	41.4	69.0	572	US-10-029-517-18	Sequence 18, Appl
4	41.4	69.0	1721	US-10-029-517-3	Sequence 3, Appl
5	41	68.3	518	US-10-029-517-101	Sequence 101, Appl
6	39.8	66.3	3343	US-10-029-517-102	Sequence 102, Appl
7	38	63.3	60	US-09-475-947A-246	Sequence 246, Appl
8	36.2	60.3	981	US-10-029-517-16	Sequence 16, Appl
9	35.6	59.3	6192	US-08-479-537A-1	Sequence 1, Appl
10	35.6	59.3	6192	US-09-083-116-1	Sequence 1, Appl
11	35.6	59.3	6192	US-09-134-916A-1	Sequence 1, Appl
12	35.6	59.3	6449	US-08-479-537A-4	Sequence 4, Appl
13	35.6	59.3	6449	US-09-083-116-4	Sequence 4, Appl
14	35.6	59.3	6449	US-09-134-916A-4	Sequence 4, Appl
15	26.6	44.3	519	US-09-646-028-42	Sequence 42, Appl
16	26.6	44.3	534	US-09-646-028-46	Sequence 46, Appl
17	25.6	42.7	392	US-09-513-999C-2024	Sequence 2024, Appl
18	24.2	40.3	38653	US-10-029-517-105	Sequence 105, Appl
19	24.2	40.3	38653	US-09-922-445-1	Sequence 1, Appl
20	23	38.3	3762	US-09-435-376-1	Sequence 1, Appl
21	22.8	38.0	1737	US-08-750-703-2	Sequence 2, Appl
22	22.8	38.0	8779	US-08-750-703-4	Sequence 4, Appl
23	22.8	38.0	77536	US-09-410-551B-1	Sequence 1, Appl
24	22.8	38.0	77536	US-09-940-316B-1	Sequence 1, Appl
25	22.8	38.0	4403765	US-09-103-840A-2	Sequence 2, Appl
26	22.8	38.0	441529	US-09-103-840A-1	Sequence 1, Appl
27	22.6	37.7	3188	US-08-943-731-183	Sequence 183, Appl

c	28	22.6	37.7	20084	3	US-08-943-731-5	Sequence 5, Appl
c	29	22.4	37.3	3900	4	US-09-023-655-1420	Sequence 1420, Appl
c	30	22.4	37.3	4403765	3	US-09-103-840A-2	Sequence 2, Appl
c	31	22.4	37.3	441529	3	US-09-103-840A-1	Sequence 1, Appl
c	32	22.2	37.0	786	4	US-09-252-991A-12436	Sequence 12436, A
c	33	22.2	37.0	1206	4	US-09-252-991A-13220	Sequence 13220, A
c	34	22.2	37.0	1719	4	US-09-252-991A-12741	Sequence 12741, A
c	35	22.2	37.0	1806	4	US-09-865-879-1	Sequence 1, Appl
c	36	22.2	37.0	2050	4	US-08-981-053-21	Sequence 21, Appl
c	37	22.2	37.0	2689	2	US-08-985-090-1	Sequence 1, Appl
c	38	22.2	37.0	2689	3	US-09-165-543-1	Sequence 1, Appl
c	39	22.2	37.0	2699	3	US-09-167-354-5	Sequence 5, Appl
c	40	22.2	37.0	2699	4	US-09-642-855-5	Sequence 5, Appl
c	41	22.2	37.0	2699	4	US-09-642-514-5	Sequence 5, Appl
c	42	22.2	37.0	47981	4	US-09-679-279-1	Sequence 1, Appl
c	43	22	36.7	8285	4	US-09-732-025-3	Sequence 3, Appl
c	44	22	36.7	11827	4	US-09-739-455-3	Sequence 3, Appl
c	45	21.8	36.3	143	4	US-09-513-999C-10197	Sequence 10197, A

ALIGNMENTS

```

RESULT 1
US-10-029-517-17
; Sequence 17, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 17
; LENGTH: 1804
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (73)...(1500)
US-10-029-517-17

Query Match      71.7%; Score 43; DB 4; Length 1804;
Best Local Similarity 83.1%; Pred. No. 4.1e-06;
Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

CY      1  GGTTCAGCGGCCCCCTGCTGACGTGTAATCGCCCGGATACGACGCGGCC 59
DB      457  GGCTCACCAGCCCCCGCCAGCCACGAGTGTGACCTGCGCCCGGACACGCGGCC 515

RESULT 2
US-10-029-517-19
; Sequence 19, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 19
; LENGTH: 8186
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 6899
; OTHER INFORMATION: unknown

```

```
/ NAME/KEY: unsure
/ LOCATION: 7155
/ OTHER INFORMATION: unknown
/ NAME/KEY: unsure
/ LOCATION: 7184
/ OTHER INFORMATION: unknown
/ NAME/KEY: unsure
/ LOCATION: 7957
/ OTHER INFORMATION: unknown
/ NAME/KEY: inttron
/ LOCATION: (2997)...(3498)
/ OTHER INFORMATION: inttron 1
/ NAME/KEY: inttron:exon junction
/ LOCATION: (3498)...(3499)
/ OTHER INFORMATION: inttron 1:exon 2
/ NAME/KEY: exon
/ LOCATION: (3508)...(3599)
/ OTHER INFORMATION: exon 2d
/ NAME/KEY: exon:inttron junction
/ LOCATION: (3982)...(3983)
/ OTHER INFORMATION: exon 2a:inttron 2a
/ NAME/KEY: inttron:exon junction
/ LOCATION: (4205)...(4206)
/ OTHER INFORMATION: inttron 2c:exon 3c
/ NAME/KEY: inttron:exon junction
/ LOCATION: (4259)...(4260)
/ OTHER INFORMATION: inttron 2d:exon 3d
/ NAME/KEY: exon
/ LOCATION: (4260)...(4328)
/ OTHER INFORMATION: exon 3d
/ NAME/KEY: inttron:exon junction
/ LOCATION: (4632)...(4633)
/ OTHER INFORMATION: inttron 3:exon 4
/ NAME/KEY: exon
/ LOCATION: (4914)...(5035)
/ OTHER INFORMATION: exon 5
/ NAME/KEY: inttron
/ LOCATION: (5265)...(6293)
/ OTHER INFORMATION: inttron 6
US-10-029-517-19
```

Query Match 71.7%; Score 43; DB 4; Length 8186;
Best Local Similarity 83.1%; Pred. No. 4.9e-06;
Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```
QY 1 GGTTCACGCGCCCCCTGCTCAGGTGTAACTCCGCCGCGATACGAGCCGCGCCC 59
DB 3825 GGCTCCACCGCCCCCGCCGCGGTGTCACTCGCCCCGAGACCAAGCCGCGCCC 3883
```

```
RESULT 3
US-10-029-517-18
/ Sequence 19, Application US/10029517
/ Patent No. 6716627
/ GENERAL INFORMATION:
/ APPLICANT: Kenneth W. Dobie
/ TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
/ FILE REFERENCE: RTS-0352
/ CURRENT APPLICATION NUMBER: US/10/029,517
/ NUMBER OF SEQ ID NOS: 107
/ SEQ ID NO 18
/ LENGTH: 572
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (67)...(572)
US-10-029-517-18
```

Query Match 69.0%; Score 41.4; DB 4; Length 572;
Best Local Similarity 81.4%; Pred. No. 1.3e-05;

```
Matches 48; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY 1 GGTTCACGCGCCCCCTGCTCAGGTGTAACTCCGCCGCGATACGAGCCGCGCCC 59
DB 478 GGCTCCACCGCCCCCGCCGCGGTGTCACTCGCCCCGAGACCAAGCCGCGCCC 536

RESULT 4
US-10-029-517-3
/ Sequence 3, Application US/10029517
/ Patent No. 6716627
/ GENERAL INFORMATION:
/ APPLICANT: Kenneth W. Dobie
/ TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
/ FILE REFERENCE: RTS-0352
/ CURRENT APPLICATION NUMBER: US/10/029,517
/ NUMBER OF SEQ ID NOS: 107
/ SEQ ID NO 3
/ LENGTH: 1721
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (58)...(1605)
US-10-029-517-3
```

Query Match 69.0%; Score 41.4; DB 4; Length 1721;
Best Local Similarity 81.4%; Pred. No. 1.5e-05;
Matches 48; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

```
QY 1 GGTTCACGCGCCCCCTGCTCAGGTGTAACTCCGCCGCGATACGAGCCGCGCCC 59
DB 442 GGCTCCACCGCCCCCGCCGCGGTGTCACTCGCCCCGAGACCAAGCCGCGCCC 500
```

```
RESULT 5
US-10-029-517-101
/ Sequence 101, Application US/10029517
/ Patent No. 6716627
/ GENERAL INFORMATION:
/ APPLICANT: Kenneth W. Dobie
/ TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
/ FILE REFERENCE: RTS-0352
/ CURRENT APPLICATION NUMBER: US/10/029,517
/ NUMBER OF SEQ ID NOS: 107
/ SEQ ID NO 101
/ LENGTH: 518
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (67)...(518)
US-10-029-517-101
```

Query Match 68.3%; Score 41; DB 4; Length 518;
Best Local Similarity 82.5%; Pred. No. 1.8e-05;
Matches 47; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```
QY 1 GGTTCACGCGCCCCCTGCTCAGGTGTAACTCCGCCGCGATACGAGCCGCGCCC 57
DB 462 GGCTCCACCGCCCCCGCCGCGGTGTCACTCGCCCCGAGACCAAGCCGCGCCC 518
```

```
RESULT 6
US-10-029-517-102
/ Sequence 102, Application US/10029517
/ Patent No. 6716627
/ GENERAL INFORMATION:
/ APPLICANT: Kenneth W. Dobie
/ TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
```


FILE REFERENCE: RTS-0352
CURRENT APPLICATION NUMBER: US/10/029,517
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 107
SEQ ID NO 102
LENGTH: 3343
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
US-10-029-517-102

Query Match 66.3%; Score 39.8; DB 4; Length 3343;
Best Local Similarity 79.7%; Pred. No. 5.7e-05;
Matches 47; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 1 GGTTCAGAGCGCCCGCTGCTCAGGTGAACATCCCGCCGAGTACAGACCGGCC 59
Db 1668 GGCTCCACCGCCCGCCCGAGCGGTGTCACTCGCGCCCGGAGACAGCGGCC 1726

RESULT 7
US-09-475-947A-246
Sequence 246, Application US/09475947A
Patent No. 6472154
GENERAL INFORMATION:
APPLICANT: Garner, Harold R.
APPLICANT: Wren, Jonathan D.
TITLE OF INVENTION: Polymorphic Repeats in Human Genes
FILE REFERENCE: UTS00667
CURRENT APPLICATION NUMBER: US/09/475,947A
CURRENT FILING DATE: 1999-12-31
NUMBER OF SEQ ID NOS: 346
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 246
LENGTH: 60
TYPE: DNA
ORGANISM: human
US-09-475-947A-246

Query Match 63.3%; Score 38; DB 4; Length 60;
Best Local Similarity 81.5%; Pred. No. 0.00015;
Matches 44; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 GGTTCAGAGCGCCCGCTGCTCAGGTGAACATCCCGCCGAGTACAGACCG 54
Db 7 GGCTCCACCGCCCGCCCGAGCGGTGTCACTCGCGCCCGGAGACAGCGGCC 60

RESULT 8
US-10-029-517-16
Sequence 16, Application US/10029517
Patent No. 6716627
GENERAL INFORMATION:
APPLICANT: Kenneth W. Dobie
APPLICANT: Susan J. Myers
TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
FILE REFERENCE: RTS-0352
CURRENT APPLICATION NUMBER: US/10/029,517
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 107
SEQ ID NO 16
LENGTH: 981
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: exon:exon junction
LOCATION: (464)..(465)
OTHER INFORMATION: exon 3b:exon 4
US-10-029-517-16

Query Match 60.3%; Score 36.2; DB 4; Length 981;
Best Local Similarity 77.2%; Pred. No. 0.00088;

Matches 44; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
Qy 1 GGTTCAGAGCGCCCGCTGCTCAGGTGAACATCCCGCCGAGTACAGACCGGCC 57
Db 21 GGCTCCACCGCCCGCCCGAGCGGTGTCACTCGCGCCCGGAGACAGCGGCC 77

RESULT 9
US-08-479-537A-1
Sequence 1, Application US/08479537A
Patent No. 5861381
GENERAL INFORMATION:
APPLICANT: CHAMBER, Pierre
APPLICANT: KIENY, Marie-Paule
APPLICANT: LATHE, Richard
APPLICANT: HAREVENT, Maïa
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
TREATMENT OR PREVENTION OF A MALIGNANT TUMOR
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/479,537A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 90/13101
FILING DATE: 23-OCT-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/FR91/00835
FILING DATE: 23-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/039,320
FILING DATE: 04-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/403,576
FILING DATE: 14-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Teskin, Robin L.
REGISTRATION NUMBER: 35,030
REFERENCE/DOCKET NUMBER: 017753-025
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 6192 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 58..120
FEATURE:
NAME/KEY: repeat region
LOCATION: 439..5239
OTHER INFORMATION: /note= "The nucleotides spanning
OTHER INFORMATION: 439-5239 constitute a repeated region wherein the repeat is 6
OTHER INFORMATION: nucleotides and encodes 20 amino acids, 17 of which are fixed
OTHER INFORMATION: The number of such repeats varies from 1 to 80."
FEATURE:
NAME/KEY: mat_peptide

LOCATION: 121..6166
FEATURE:
NAME/KEY: repeat_region
LOCATION: 457
OTHER INFORMATION: /note= "Nucleotide 457 is X1 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 487
OTHER INFORMATION: /note= "Nucleotide 487 is Y = NNN
OTHER INFORMATION: which is the codon for Thr or Asn wherein Thr = ACT, ACC, ACA
OTHER INFORMATION: or ACG; and Asn = AAT or AAC."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 496
OTHER INFORMATION: /note= "Nucleotide 496 is X2 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
US-08-479-537A-1

Query Match 59.3%; Score 35.6; DB 2; Length 6192;
Best Local Similarity 69.5%; Pred. No. 0.0018;
Matches 41; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1 GGTTCACGCGCCCTCTGCTCAGCGTGTAAACATCCGCCGAGATACGACGCGGCC 59
DB 442 GGCTCCACGCGCCCTCTGCTCAGCGTGTAAACATCCGCCGAGATACGACGCGGCC 500

RESULT 10
US-09-083-116-1
Sequence 1, Application US/09083116
Patent No. 6203795

GENERAL INFORMATION:
APPLICANT: CHAMON, Pierre
APPLICANT: KIENY, Marie-Paule
APPLICANT: LATHE, Richard
APPLICANT: HAREUVENI, Mara
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESSES:
ADDRESS: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/083,116
FILING DATE:

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/479,537
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/FR91/00835
FILING DATE: 23-OCT-1991
APPLICATION DATA:
APPLICATION NUMBER: US 08/039,320
FILING DATE: 04-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/403,576
FILING DATE: 14-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Teskin, Robin L.
REGISTRATION NUMBER: 35,030

REFERENCE/DOCKET NUMBER: 017753-025
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 6192 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: sig peptide
LOCATION: 58..120
FEATURE:
NAME/KEY: repeat_region
LOCATION: 439..5239
OTHER INFORMATION: /note= "The nucleotides spanning
OTHER INFORMATION: 439-5239 constitute a repeated region wherein the repeat is 6
OTHER INFORMATION: nucleotides and encodes 20 amino acids, 17 of which are fixe
OTHER INFORMATION: the number of such repeats varies from 1 to 80."
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 121..6166
FEATURE:
NAME/KEY: repeat_region
LOCATION: 457
OTHER INFORMATION: /note= "Nucleotide 457 is X1 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CC
OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 487
OTHER INFORMATION: /note= "Nucleotide 487 is Y = NNN
OTHER INFORMATION: which is the codon for Thr or Asn wherein Thr = ACT, ACC, ACA
OTHER INFORMATION: or ACG; and Asn = AAT or AAC."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 496
OTHER INFORMATION: /note= "Nucleotide 496 is X2 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
US-09-083-116-1

Query Match 59.3%; Score 35.6; DB 3; Length 6192;
Best Local Similarity 69.5%; Pred. No. 0.0018;
Matches 41; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1 GGTTCACGCGCCCTCTGCTCAGCGTGTAAACATCCGCCGAGATACGACGCGGCC 59
DB 442 GGCTCCACGCGCCCTCTGCTCAGCGTGTAAACATCCGCCGAGATACGACGCGGCC 500

RESULT 11
US-09-134-916A-1
Sequence 1, Application US/09134916A
Patent No. 6328956

GENERAL INFORMATION:
APPLICANT: CHAMON, Pierre
APPLICANT: KIENY, Marie-Paule
APPLICANT: LATHE, Richard
APPLICANT: HAREUVENI, Mara
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESSES:
ADDRESS: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/134,916A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/479,537
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: FR 90/13101
FILING DATE: 23-OCT-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/FR91/00835
FILING DATE: 23-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/039,320
FILING DATE: 04-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/403,576
FILING DATE: 14-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Teskin, Robin L.
REGISTRATION NUMBER: 35,030
REFERENCE/DOCKET NUMBER: 017753-025
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-6620
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 6192 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 58..120
FEATURE:
NAME/KEY: repeat_region
LOCATION: 439..5239
OTHER INFORMATION: /note= "The nucleotides spanning 439-5239 constitute a repeated region wherein the repeat is 6
OTHER INFORMATION: nucleotides and encodes 20 amino acids, 17 of which are fixed
OTHER INFORMATION: The number of such repeats varies from 1 to 80."
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 121..6166
FEATURE:
NAME/KEY: repeat_region
LOCATION: 457
OTHER INFORMATION: /note= "Nucleotide 457 is X1 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
OTHER INFORMATION: or CCG, and Ala = GCT, GCC, GCA, or GCG."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 487
OTHER INFORMATION: /note= "Nucleotide 487 is Y = NNN
OTHER INFORMATION: which is the codon for Thr or Asn wherein Thr = ACT, ACC, ACA
OTHER INFORMATION: or ACG, and Asn = AAT or AAC."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 496
OTHER INFORMATION: /note= "Nucleotide 496 is X2 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
OTHER INFORMATION: or CCG, and Ala = GCT, GCC, GCA, or GCG."
US-09-134-916A-1
Query Match 59.3%; Score 35.6; DB 3; Length 6192;
Best Local Similarity 69.5%; Pred. No. 0.0018;
Matches 41; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1 GGTCGACGCGCCCCCTGCTCATCGGTGAACATCCGCGGAGTACGACCGCGCCC 59
DB 442 GGCTCAGCGCGCCCGCCNNNCCGACGCTGTCACCTCGCGCCCGGACNNNAGCGCGNNCC 500
RESULT 12
US-08-479-537A-4
Sequence 4, Application US/08479537A
Patent No. 5861381
GENERAL INFORMATION:
APPLICANT: CHAMBERN, Pierre
APPLICANT: KIENY, Marie-Paule
APPLICANT: LATHE, Richard
APPLICANT: HAREIVENT, Mara
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
TREATMENT OR PREVENTION OF A MALIGNANT TUMOR
TITLE OF INVENTION: 5
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESSES:
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/479,537A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 90/13101
FILING DATE: 23-OCT-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/FR91/00835
FILING DATE: 23-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/039,320
FILING DATE: 04-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/403,576
FILING DATE: 14-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Teskin, Robin L.
REGISTRATION NUMBER: 35,030
REFERENCE/DOCKET NUMBER: 017753-025
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-6620
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 6449 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 58..120
FEATURE:
NAME/KEY: repeat_region
LOCATION: 439..5239
OTHER INFORMATION: /note= "The nucleotides spanning 439-5239 constitute a repeated region wherein the repeat is 6
OTHER INFORMATION: nucleotides and encodes 20 amino acids, 17 of which are fixed
OTHER INFORMATION: The number of such repeats varies from 1 to 80."
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 121..5661
FEATURE:

```

NAME/KEY: repeat_region
LOCATION: 457
OTHER INFORMATION: /note= "Nucleotide 457 is X1 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 487
OTHER INFORMATION: /note= "Nucleotide 487 is Y = NNN
OTHER INFORMATION: which is the codon for Thr or Asn wherein Thr = ACT, ACC, ACA
OTHER INFORMATION: or ACG; and Asn = AAT or AAC."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 496
OTHER INFORMATION: /note= "Nucleotide 496 is X2 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
US-08-479-537A-4

Query Match          59.3%; Score 35.6; DB 2; Length 6449;
Best Local Similarity 69.5%; Pred. No. 0.0018;
Matches 41; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1 GGTTGACGCGCCCCCTGCTCAGGTGTACATCCGCGCCGATACGACCGGCCCC 59
Db 442 GGCTCCACCGCCCCCNNGCCACGAGTGTACCTCGGCGCCGACNNNAGCCGNNNCC 500

RESULT 13
US-09-083-116-4
Sequence 4, Application US/09083116
Patent No. 6203795
GENERAL INFORMATION:
APPLICANT: CHAMRON, Pierre
APPLICANT: KIENY, Marie-Paule
APPLICANT: LATHE, Richard
APPLICANT: HAREUVENI, Mara
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
TITLE OF INVENTION: TREATMENT OR PREVENTION OF A MALIGNANT TUMOR
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/083,116
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/479,537
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/FR91/00835
FILING DATE: 23-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/039,320
FILING DATE: 04-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/403,576
FILING DATE: 14-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Teskin, Robin L.
REGISTRATION NUMBER: 35,030
REFERENCE/DOCKET NUMBER: 017753-025
TELECOMMUNICATION INFORMATION:
```

```

TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 6449 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 58..120
FEATURE:
NAME/KEY: repeat_region
LOCATION: 439..5239
OTHER INFORMATION: /note= "The nucleotides spanning
OTHER INFORMATION: 439-5239 constitute a repeated region wherein the repeat is 6
OTHER INFORMATION: nucleotides and encodes 20 amino acids, 17 of which are fixed
OTHER INFORMATION: The number of such repeats varies from 1 to 80."
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 121..5661
FEATURE:
NAME/KEY: repeat_region
LOCATION: 457
OTHER INFORMATION: /note= "Nucleotide 457 is X1 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 487
OTHER INFORMATION: /note= "Nucleotide 487 is Y = NNN
OTHER INFORMATION: which is the codon for Thr or Asn wherein Thr = ACT, ACC, ACA
OTHER INFORMATION: or ACG; and Asn = AAT or AAC."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 496
OTHER INFORMATION: /note= "Nucleotide 496 is X2 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
US-09-083-116-4

Query Match          59.3%; Score 35.6; DB 3; Length 6449;
Best Local Similarity 69.5%; Pred. No. 0.0018;
Matches 41; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1 GGTTGACGCGCCCCCTGCTCAGGTGTACATCCGCGCCGATACGACCGGCCCC 59
Db 442 GGCTCCACCGCCCCCNNGCCACGAGTGTACCTCGGCGCCGACNNNAGCCGNNNCC 500

RESULT 14
US-09-134-916A-4
Sequence 4, Application US/09134916A
Patent No. 6328956
GENERAL INFORMATION:
APPLICANT: CHAMRON, Pierre
APPLICANT: KIENY, Marie-Paule
APPLICANT: LATHE, Richard
APPLICANT: HAREUVENI, Mara
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
TITLE OF INVENTION: TREATMENT OR PREVENTION OF A MALIGNANT TUMOR
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
```

```

1 OPERATING SYSTEM: PC-DOS/MS-DOS
2 SOFTWARE: Patent In Release #1.0, Version #1.30
3 CURRENT APPLICATION DATA:
4 APPLICATION NUMBER: US/09/134,916A
5 FILING DATE:
6 CLASSIFICATION:
7 PRIOR APPLICATION DATA:
8 APPLICATION NUMBER: US/08/479,537
9 FILING DATE: 07-JUN-1995
10 APPLICATION NUMBER: FR 90/13101
11 FILING DATE: 23-OCT-1990
12 PRIOR APPLICATION DATA:
13 APPLICATION NUMBER: WO PCT/FR91/00835
14 FILING DATE: 23-OCT-1991
15 PRIOR APPLICATION DATA:
16 APPLICATION NUMBER: US 08/039,320
17 FILING DATE: 04-APR-1993
18 PRIOR APPLICATION DATA:
19 APPLICATION NUMBER: US 08/403,576
20 FILING DATE: 14-MAR-1995
21 ATTORNEY/AGENT INFORMATION:
22 NAME: Teskin, Robin L.
23 REGISTRATION NUMBER: 35,030
24 REFERENCE/DOCKET NUMBER: 017753-025
25 TELECOMMUNICATION INFORMATION:
26 TELEPHONE: (703) 836-6620
27 TELEFAX: (703) 836-6021
28 INFORMATION FOR SEQ ID NO: 4:
29 SEQUENCE CHARACTERISTICS:
30 LENGTH: 6449 base pairs
31 TYPE: nucleic acid
32 STRANDEDNESS: single
33 TOPOLOGY: linear
34 MOLECULE TYPE: DNA (genomic)
35 FEATURE:
36 NAME/KEY: sig_peptide
37 LOCATION: 58..120
38 FEATURE:
39 NAME/KEY: repeat_region
40 LOCATION: 439..5339
41 OTHER INFORMATION: /note= "The nucleotides spanning
42 OTHER INFORMATION: 439-5239 constitute a repeated region wherein the repeat is 6
43 OTHER INFORMATION: nucleotides and encodes 20 amino acids, 17 of which are fixed
44 OTHER INFORMATION: The number of such repeats varies from 1 to 80."
45 FEATURE:
46 NAME/KEY: mat_peptide
47 LOCATION: 121..5661
48 FEATURE:
49 NAME/KEY: repeat_region
50 LOCATION: 457
51 OTHER INFORMATION: /note= "Nucleotide 457 is X1 = NNN
52 OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
53 OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
54 FEATURE:
55 NAME/KEY: repeat_region
56 LOCATION: 487
57 OTHER INFORMATION: /note= "Nucleotide 487 is Y = NNN
58 OTHER INFORMATION: which is the codon for Thr or Asn wherein Thr = ACT, ACC, ACG
59 OTHER INFORMATION: or ACG; and Asn = AAT or AAC."
60 FEATURE:
61 NAME/KEY: repeat_region
62 LOCATION: 496
63 OTHER INFORMATION: /note= "Nucleotide 496 is X2 = NNN
64 OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
65 OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
87
```

```

DB
442 GAGCTCAGCCGCCCCNNNGCCCAAGGTGTACCTCGGAGCCCGGACNNNAGGCGCANNCC
500

RESULT 15
US-09-646-028-42
; Sequence 42, Application US/09646028
; Patent No. 6562347
; GENERAL INFORMATION:
; APPLICANT: Kwak, Larry
; APPLICANT: Biragyn, Arva
; TITLE OF INVENTION: METHODS AND COMPOSITIONS OF
; FILE REFERENCE: 14014.0316/P
; CURRENT APPLICATION NUMBER: US/09/646, 028
; CURRENT FILING DATE: 2000-09-12
; PRIOR APPLICATION NUMBER: 60/077, 745
; FILING DATE: 1998-03-12
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 42
; LENGTH: 519
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of artificial sequence: /note=synthetic construct
US-09-646-028-42

```

```

Query Match          44.3%: Score 26.6; DB 4, Length 519;
Best Local Similarity 78.0%: Pred. No. 1.7;
Matches 32; Conservative 0; Mismatches 9; Indels 0; Gaps 0.

QY      1  GATTGACGCGCCCGCTGCACAGGTATACATCCGCCCC 41
        |||||  |||||  |||||  |||||  |||||  |||||
Db       331 GATTCTACTGCTCCGCGGCACAGGTATACCTTCTGCCGC 371

Search completed: January 16, 2005, 03:01:37
Job time : 44.2 secs

```

by 1 GGTTCAGACGGCCCCCTGCTCAGCGGTGAACATCCGCCCCGGATACAGACCGGCCCC 59

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2005, 19:43:55 : Search time 1550.6 Seconds
(without alignments)
1410.024 Million cell updates/sec

Title: US-10-057-136-7

Perfect score: 60

Sequence: 1 ggttcagacgccccctgc.....cggatcacgacgccccct 60

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

EST :
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_g881:*
9: gb_g882:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	43	71.7	330	1	AI925867 WC2004.x
2	43	71.7	604	4	BM791359 K-EST0071
3	43	71.7	877	5	BU542454 AGENCOURT
4	43	71.7	959	6	CA489836 AGENCOURT
5	43	71.7	1113	5	BU148487 AGENCOURT
6	43	71.7	1130	5	BU542996 AGENCOURT
7	43	71.7	1234	5	BQ936898 AGENCOURT
8	43	71.7	1268	5	BQ943554 AGENCOURT
9	43	71.7	1343	5	BQ920055 AGENCOURT
10	43	71.7	1349	5	BU152566 AGENCOURT
11	43	71.7	1420	5	BU542790 AGENCOURT
12	43	71.7	1531	5	BU543309 AGENCOURT
13	41.4	69.0	1262	5	BQ935496 AGENCOURT
14	36.8	61.3	1334	5	BQ943809 AGENCOURT
15	36.2	60.3	981	4	BG774910 602649832
16	36.2	60.3	1536	5	BQ923149 AGENCOURT
17	35.2	58.7	166	7	T27692
18	35.2	58.7	475	6	CB120860 K-EST0168
19	35.2	58.7	669	4	CB122585 K-EST0170
20	33.8	56.3	619	4	B1260921 602970962
21	33.4	55.7	472	4	BM759495 K-EST0039
22	28.6	47.7	878	2	BF786059 602110629
23	28.6	47.7	878	2	BF786059 602110629
24	28.6	47.7	878	2	CA924184 MTU7CL.P2

25	28	46.7	635	6	CA924567 MTU7CL.P7
26	27.2	45.3	595	2	BF825165 RC3-HN002
27	27.2	45.3	504	1	AJ280697 4A3A-AAS-
28	27	45.0	811	4	BI956680 HYSMH000
29	27	45.0	886	7	CL653412 PRI0118D-
30	26.6	44.3	308	7	BM427484 P9f2n.pko
31	26.6	44.3	478	4	BI518647 603062003
32	26.6	44.3	832	4	BM337143 BX337143
33	26.6	44.3	750	5	BI523776 603052074
34	26.4	44.0	984	4	BI523854 603051982
35	26.2	43.7	1868	4	BG336063 602404834
36	26.2	43.7	548	2	BE706360 RCI-H2025
37	26.2	43.3	607	9	CE187066 t3gr-g8s-
38	26	43.3	682	9	AG082656 Pan_trog1
39	26	43.3	1119	4	BM911965 AGENCOURT
40	25.6	42.7	346	5	BY098559 BY098559
41	25.6	42.7	435	8	BH253383 SALK_0148
42	25.6	42.7	467	9	CC719739 OCKAD34TH
43	25.6	42.7	549	7	CF482287 POL1_6 CO
44	25.6	42.7	549	7	CF482287 POL1_6 CO
45	25.6	42.7	653	9	CC623364 OCUT190TH

ALIGNMENTS

RESULT 1
AI925867/c 330 bp mRNA linear EST 08-MAR-2000
LOCUS WO2004.x1 NCI CGAP Paul Homo sapiens CDNA clone IMAGE:2455879 3'
DEFINITION similar to contains element MSRI MSRI repetitive element ;, mRNA
sequence.

ACCESSION AI925867
VERSION AI925867.1 GI:5661831
KEYWORDS EST.
SOURCE Homo sapiens
ORGANISM Homo sapiens (human)

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 330)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov
Life Technologies catalog #: 11548-013
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/HLN at:
www.bio.lml.gov/bbip/image/image.html
Insert length: 578 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence strop: 305.

FEATURES
source location/Qualifiers
1..330
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2455879"
/issue_type="adenocarcinoma"
/lab_host="DH10B"
/clone_lib="NCI-CGAP_Paul"
/note="Organ: pancreas; Vector: pCMV-SPORT6; Site: 1; Salt; Site: 2; Ncti; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.72 kb. Life Technologies catalog #: 11548-013"

ORIGIN

Query Match 71.7%; Score 43; DB 1; Length 330;
Best Local Similarity 83.1%; Pred. No. 0.00058;
Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 GGTTTCAGCGGCCCCCTGCTCAAGGTGTAACTCCGCCCCGGATACCAAGACCGGCCCC 59
 Db 316 GGCTTCACGCGCCCCCGCCACGCGGTGTCACTCGGCCCCCGAGACACGAGCGGCCCC 258

RESULT 2
 LOCUS BM791359
 DEFINITION K-EST0071342 S21SNUS20 Homo sapiens cDNA clone S21SNUS20-14-A06 5', mRNA sequence.
 ACCESSION BM791359
 VERSION BM791359.1 GI:19139591
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 604)
 Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R., Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and Kim,Y.S.
 21C Frontier Korean EST Project 2001
 JOURNAL Unpublished (2002)
 COMMENT Contact: Kim YS
 Genome Research Center
 Korea Research Institute of Bioscience & Biotechnology
 52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
 Tel: +82-42-860-4470
 Fax: +82-42-860-4409
 Email: yongsung@mail.krrib.re.kr
 Plate: 14 row: A column: 06
 High quality sequence stop: 604.
 Location/Qualifiers
 1..604
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="S21SNUS20-14-A06"
 /sex="F"
 /tissue_type="Stomach"
 /cell_type="floating aggregates"
 /cell_line="SNU-520"
 /lab_host="Top10F"
 /clone_id="S21SNUS20"
 /note="Organ: Stomach; Vector: pTZ18RP1; Site 1: EcoRI; Site 2: NotI; The poly (A)+ RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then decapped with tobacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA-RNA linker including EcoR I site by treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dt-selected mRNA by priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand obtained cDNA vectors were used for transformation of competent cells E. coli Top10F by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

ORIGIN

Query Match 71.7%; Score 43; DB 4; Length 604;
 Best Local Similarity 83.1%; Pred. No. 0.00059;
 Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 GGTTTCAGCGGCCCCCTGCTCAAGGTGTAACTCCGCCCCGGATACCAAGACCGGCCCC 59
 Db 106 GGCTTCACGCGCCCCCGCCACGCGGTGTCACTCGGCCCCCGAGACACGAGCGGCCCC 164

RESULT 3
 BU542454

LOCUS BU542454 877 bp mRNA linear EST 13-SEP-2002
 DEFINITION AGENCOURT_10322173 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6574322 5', mRNA sequence.
 ACCESSION BU542454
 VERSION BU542454.1 GI:22852937
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 877)
 NIH-MGC http://mgc.nci.nih.gov/.
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
 COMMENT Unpublished (1999)
 Contact: Robert Strusberg, Ph.D.
 Email: cgepds@mail.nih.gov
 Tissue Procurement: DCTD/DTF
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
 http://image.lnl.gov
 Plate: LNCM2769 row: h column: 02
 High quality sequence stop: 760.
 Location/Qualifiers
 1..877
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6574322"
 /tissue_type="carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_id="NIH MGC 40"
 /note="Organ: prostate; Vector: pORF7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dt priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGAG(G). Library constructed by Ling Hong in the Laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC library."

ORIGIN

Query Match 71.7%; Score 43; DB 5; Length 877;
 Best Local Similarity 83.1%; Pred. No. 0.00059;
 Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 GGTTTCAGCGGCCCCCTGCTCAAGGTGTAACTCCGCCCCGGATACCAAGACCGGCCCC 59
 Db 35 GGCTTCACGCGCCCCCGCCACGCGGTGTCACTCGGCCCCCGAGACACGAGCGGCCCC 93

RESULT 4

CA489836

LOCUS CA489836 959 bp mRNA linear EST 14-NOV-2002
 DEFINITION AGENCOURT_10810668 MAPcL Homo sapiens cDNA clone IMAGE:672324 5', mRNA sequence.
 ACCESSION CA489836
 VERSION CA489836.1 GI:24952627
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 959)
 NIH-MGC http://mgc.nci.nih.gov/.
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
 COMMENT Unpublished (1999)
 Contact: Robert Strusberg, Ph.D.
 Email: cgepds@mail.nih.gov
 Tissue Procurement: Kisti A. Eglund, Ira Pastan
 cDNA Library Preparation: Invitrogen Corp

REFERENCE 1 (bases 1 to 959)
 NIH-MGC http://mgc.nci.nih.gov/.
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
 COMMENT Unpublished (1999)
 Contact: Robert Strusberg, Ph.D.
 Email: cgepds@mail.nih.gov
 Tissue Procurement: Kisti A. Eglund, Ira Pastan
 cDNA Library Preparation: Invitrogen Corp

CDNA library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLNL4284 row: n column: 20
High quality sequence start: 31
High quality sequence stop: 446.
Location/Qualifiers

ORIGIN

Query Match	71.7%	Score	43	DB	6	Length	959
Best Local Similarity	83.1%	Pred. No.	0.00659				
Matches	49	Conservative	0	Mismatches	10	Indels	0
						Gaps	0

Qy 1 GATTGACGGCCCCCTGCTCACGGTGTAAATCCGCCCCGGATACAGACCGCCCC 59
Db 349 GGCTCAACGCCCCCAGCCACGAGTGTCACTCGCCCCCGGACACAGAGCGGCCCC 400

RESULT 5	1113 bp	mbNA	linear	EST 03-SEP-2002
BU148487				
LOCUS				
DEFINITION	BU148487	AGENCOURT.8670479 NIH_MGC_40 Homo sapiens	cdna clone	IMAGE:6380642
ACCESSION	BU148487	5', mRNA sequence.		
VERSION	BU148487.1	GI:22626219		
KEYWORDS	EST.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
	Eumetazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
REFERENCE	1 (bases 1 to 1113)			
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .			
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)			
JOURNAL	Unpublished (1995)			
COMMENT	Contact: Robert Strausberg, Ph.D.			

FEATURES
SOURCE

```
/organism="Homo sapiens"  
/mol_type="mRNA"  
/db_xref="taxon:9606"  
/clone="IMAGE:5380642"  
/tissue_type="carcinoma, cell line"  
/lab_host="DH10B (phage-resistant)"  
/clone_lib="NH_MGC_40"
```

/site2="Organ: prostate; Vector: DOT3; Site 1: XhoI;
 Site 2: EcoRI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCACGAG(G). Library constructed by
 Ling Hong in the laboratory of Gerald M. Rubin (University
 of California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library."

KEYWORDS

ORGANISM	Homo sapiens
Eukaryota; Metazoa; Chordata; Mammalia; Eutheria; Primates; Hominidae	

AUTHORS NIH-MGC <http://mgc.ncl.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Unpublished (1999)

Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC Clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: L1CM2771 row: c column: 11
High quality sequence start: 27
High quality sequence stop: 246.

ORIGIN

Query Match	71.74;	Score 43;	DB 5;	Length 1130;
Best Local Similarity	83.18;	Pred. No. 0.0006;		
Matches 49;	Conservative 0;	Mismatches 10;	Indels 0;	Gaps 0;

Qy 1 GGTTCACAGGCCCCCCCCGTCTACAGGTATACATCCGGCCCCGGATACCAAGCCGGCCCC 59
Ddb 69 GGCTCACAGGCCCCCCCCAGCCACAGGTGTCACTTCGGCCCCGGACACCAAGCCGGCCCC 127

RESULT 7
BO936898 1234 bp mRNA linear EST 21-AUG-2002
LOCUS AGENCOURT_9919110 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6383855
DEFINITION 5', mRNA sequence.
ACCESSION BO936898
VERSION BO936898.1 GI:22352281
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1234)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LNCM2577 row: 0 column: 24
High quality sequence stop: 245.
Location/Qualifiers
1..1234
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6383855"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
/note="Organ: prostate; Vector: pOT7; Site: 1: XhoI;
Site 2: EcoRI; CDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGACGAG(G). Library constructed by
ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-CDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN
Query Match 71.7%; Score 43; DB 5; Length 1234;
Best Local Similarity 83.1%; Pred. No. 0.0006;
Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
Qy 1 GGTTCAGCGGCCCCCTGTCTCAGGTGAACATCCGCCGATACAGACCGGCCCC 59
Db 30 GGCTCCACCGCCCCCGACGCCGATGTCACCTGCGCCCGACACGCGCGCCCC 86

RESULT 8
BO943554 1268 bp mRNA linear EST 21-AUG-2002
LOCUS AGENCOURT_877815 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6384308
DEFINITION 5', mRNA sequence.
ACCESSION BO943554
VERSION BO943554.1 GI:22359032
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1268)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)

COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LNCM2579 row: b column: 21
High quality sequence stop: 177.
Location/Qualifiers
1..1268
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6384308"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
/note="Organ: prostate; Vector: pOT7; Site: 1: XhoI;
Site 2: EcoRI; CDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGACGAG(G). Library constructed by
ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-CDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN
Query Match 71.7%; Score 43; DB 5; Length 1268;
Best Local Similarity 83.1%; Pred. No. 0.0006;
Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
Qy 1 GGTTCAGCGGCCCCCTGTCTCAGGTGAACATCCGCCGATACAGACCGGCCCC 59
Db 90 GGCTCCACCGCCCCCGACGCCGATGTCACCTGCGCCCGACACGCGCGCCCC 148

RESULT 9
BO920055 1343 bp mRNA linear EST 20-AUG-2002
LOCUS AGENCOURT_10031674 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6481609
DEFINITION 5', mRNA sequence.
ACCESSION BO920055
VERSION BO920055.1 GI:2234753
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1343)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LNCM2663 row: e column: 02
High quality sequence start: 56
High quality sequence stop: 237.
Location/Qualifiers
1..1343
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6481609"

/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_1lb="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC library."

ORIGIN

Query Match 71.7%; Score 43; DB 5; Length 1349;
Best Local Similarity 83.1%; Pred. No. 0.0006;
Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 GGTTCAGCGCCCCCTCTCAGGTGTAACTCCGCCGATACGACCGGCCCC 59
|||
72 GGCTCAGCGCCCCCGCCAGCGGTGTCACTCTGCCCGACACGAGCGGCCCC 130
|||

RESULT 10 1349 bp mRNA linear EST 03-SEP-2002
LOCUS BU152566
DEFINITION AGENCOURT_8670683 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6380559
5', mRNA sequence.
ACCESSION BU152566
VERSION BU152566.1 GI:22666098
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 1349)
NIH-MGC http://mhc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: DCTD/DRP
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2569 row: f column: 16
High quality sequence stop: 291.

FEATURES

source

1.1349
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_1lb="IMAGE:6380559"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_1lb="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC library."

ORIGIN

Query Match 71.7%; Score 43; DB 5; Length 1349;
Best Local Similarity 83.1%; Pred. No. 0.0006;
Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 GGTTCAGCGCCCCCTCTCAGGTGTAACTCCGCCGATACGACCGGCCCC 59
|||

Db 19 GGCTCAGCGCCCCCGCCAGCGGTGTCACTCTGCCCGACACGAGCGGCCCC 77
|||

RESULT 11 1420 bp mRNA linear EST 13-SEP-2002
LOCUS BU542790
DEFINITION AGENCOURT_10334841 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6574725
5', mRNA sequence.
ACCESSION BU542790
VERSION BU542790.1 GI:22853273
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 1420)
NIH-MGC http://mhc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: DCTD/DRP
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2770 row: h column: 21
High quality sequence stop: 288.

FEATURES

source

1.1420
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_1lb="IMAGE:6574725"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_1lb="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC library."

ORIGIN

Query Match 71.7%; Score 43; DB 5; Length 1420;
Best Local Similarity 83.1%; Pred. No. 0.0006;
Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 GGTTCAGCGCCCCCTCTCAGGTGTAACTCCGCCGATACGACCGGCCCC 59
|||
19 GGCTCAGCGCCCCCGCCAGCGGTGTCACTCTGCCCGACACGAGCGGCCCC 77
|||

RESULT 12 1531 bp mRNA linear EST 13-SEP-2002
LOCUS BU543309
DEFINITION AGENCOURT_10327072 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6575345
5', mRNA sequence.
ACCESSION BU543309
VERSION BU543309.1 GI:22853792
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 1531)
NIH-MGC http://mhc.nci.nih.gov/.

TITLE
JOURNAL
COMMENT
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LNCM2772 row: b column: 21
High quality sequence stop: 166.

FEATURES
source
Location/Qualifiers
1..1531
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6575349"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCAAGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN
Query Match 71.7%; Score 43; DB 5; Length 1531;
Best Local Similarity 83.1%; Pred. No. 0.0006;
Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
Qy 1 GGTTGACGGCCCCCTGCTCAGGTGTAATCGCCCGGATACGACGGCCCC 59
Db 30 GGCTCCACGGCCCCCGCCAGCCAGGTGTCACTCTGCCCCGACACGCGCCCC 88

RESULT 13
LOCUS BQ935496 1262 bp mRNA linear EST 21-AUG-2002
DEFINITION AGENCOURT 8776475 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6382088
5' mRNA sequence.
ACCESSION BQ935496
VERSION BQ935496.1 GI:22350879
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1262)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LNCM2573 row: f column: 09
High quality sequence start: 46
High quality sequence stop: 157.

FEATURES
source
Location/Qualifiers
1..1262
/organism="Homo sapiens"
/mol_type="mRNA"

/db_xref="taxon:9606"
/clone="IMAGE:6382088"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCAAGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN
Query Match 69.0%; Score 41.4; DB 5; Length 1262;
Best Local Similarity 81.4%; Pred. No. 0.002;
Matches 48; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
Qy 1 GGTTGACGGCCCCCTGCTCAGGTGTAATCGCCCGGATACGACGGCCCC 59
Db 56 GGCTCCACGGCCCCCGCCAGCCAGGTGTCACTCTGCCCCGACACGCGCCCC 114

RESULT 14
LOCUS BQ943809 1334 bp mRNA linear EST 21-AUG-2002
DEFINITION AGENCOURT 8775676 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6383286
5' mRNA sequence.
ACCESSION BQ943809
VERSION BQ943809.1 GI:22359287
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1334)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LNCM2576 row: h column: 07
High quality sequence stop: 114.

FEATURES
source
Location/Qualifiers
1..1334
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6383286"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCAAGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN
Query Match 61.3%; Score 36.8; DB 5; Length 1334;
Best Local Similarity 85.4%; Pred. No. 0.07;
Matches 41; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 12 CCCCCCTGCTCAGGATGTAACATCCGCGGATACAGACCGCCCC 59
 |||||
 Db 43 CCCCCAGCCCAAGGTGTCACTCGGCGGACACCAAGCGGCCCC 90

RESULT 15

BG774910

602649832F1 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:4761054 5',

mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 981)
 NIH-MGC <http://mgs.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-r@mail.nih.gov
 Tissue Procurement: DCTD/DRP

cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLCM1612 row: 0 column: 07
 High quality sequence stop: 874.
 Location/Qualifiers

FEATURES

source

1..981
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4761054"
 /tissue_type="carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 40"
 /note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
 Site_2: EcoRI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCACGAG(G). Library constructed by
 Ling Hong in the laboratory of Gerald M. Rubin (University
 of California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library."

ORIGIN

Query Match 60.3%; Score 36.2; DB 4; Length 981;

Best Local Similarity 77.2%; Pred. No. 0.11; Mismatches 13; Indels 0; Gaps 0;

Matches 44; Conservative 0; Db 1 GGTTCACGCGCCCCCTGCTCAGGATGTAACATCCGCGGATACAGACCGGCC 57
 |||||
 21 GGCTTCACGCGCGCCCCCAGCCATGCTGTCACTCGGCGGACACCAAGCGGCCCC 77

Search completed: January 16, 2005, 02:55:30
 Job time : 1552.6 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2005, 16:00:11 ; Search time 170.5 Seconds
(without alignments)
1847.304 Million cell updates/sec

Title: US-10-057-136-7

Perfect score: 60

Sequence: 1 ggttcgacggccccccctgc.....cggataccgacggccct 60

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseg 23Sep04:*

1: genesegn1980s:*
2: genesegn1990s:*
3: genesegn2000s:*
4: genesegn2001as:*
5: genesegn2001bs:*
6: genesegn2002as:*
7: genesegn2002bs:*
8: genesegn2003as:*
9: genesegn2003bs:*
10: genesegn2003cs:*
11: genesegn2003ds:*
12: genesegn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	60	2 AAV48320	AAV48320 Nucleotide
2	43	71.7	60	2 AAV48316	AAV48316 Nucleotide
3	43	71.7	309	1 AAN90579	Aan90579 pDF9.3 CD
4	43	71.7	1194	12 AD157712	Ad157712 Human bre
5	43	71.7	1378	12 AD157693	Ad157693 Human bre
6	43	71.7	1424	12 ADO23180	ADO23180 Antisense
7	43	71.7	1428	6 ABL60159	Ab160159 Human MUC
8	43	71.7	1428	12 ADO23125	ADO23125 Human MUC
9	43	71.7	1457	12 ADF32627	Adf32627 Plaemid J
10	43	71.7	1527	2 AAV48329	AAV48329 MiniMUC1
11	43	71.7	1614	12 ADR70370	Ad70370 Respirato
12	43	71.7	1630	12 AD157708	Ad157708 Human bre
13	43	71.7	1634	12 AD157689	Ad157689 Human bre
14	43	71.7	1712	12 AD157686	Ad157686 Human bre
15	43	71.7	1738	12 AD157669	Ad157669 Human bre
16	43	71.7	1755	12 AD157673	Ad157673 Human bre
17	43	71.7	1774	12 ADE43591	Ad43591 Plaemid J
18	43	71.7	1774	12 ADF32625	Adf32625 Plaemid J
19	43	71.7	1799	12 ADO23124	ADO23124 Human MUC
20	43	71.7	1803	12 AD157699	Ad157699 Human bre
21	43	71.7	1804	6 ABL67539	Ab167539 Thyroid c

22	43	71.7	1804	9 AAD56950	Aad56950 Human muc
23	43	71.7	1804	10 ADD14719	Add14719 Human src
24	43	71.7	1804	12 ADP13294	Adp13294 Renal cel
25	43	71.7	1805	12 ADO28642	Ado28642 Human MUC
26	43	71.7	1808	12 AD157706	Ad157706 Human bre
27	43	71.7	1818	12 ADF32633	Adf32633 Plaemid J
28	43	71.7	1823	6 ABZ35228	Abz35228 Human gen
29	43	71.7	1823	12 AD157707	Ad157707 Human bre
30	43	71.7	1835	12 ADF32631	Adf32631 Plaemid J
31	43	71.7	1874	12 AD157688	Ad157688 Human bre
32	43	71.7	1882	12 AD157677	Ad157677 Human bre
33	43	71.7	1918	12 AD157672	Ad157672 Human bre
34	43	71.7	1930	12 AD157678	Ad157678 Human bre
35	43	71.7	1945	12 AD157676	Ad157676 Human bre
36	43	71.7	1949	12 AD157698	Ad157698 Human bre
37	43	71.7	1953	12 AD157668	Ad157668 Human bre
38	43	71.7	2045	12 AD157701	Ad157701 Human bre
39	43	71.7	2049	12 AD157682	Ad157682 Human bre
40	43	71.7	2090	12 AD157705	Ad157705 Human bre
41	43	71.7	2094	12 AD157681	Ad157681 Human bre
42	43	71.7	2135	12 ADF32629	Adf32629 Plaemid J
43	43	71.7	2194	12 AD157683	Ad157683 Human bre
44	43	71.7	2255	12 AD157667	Ad157667 Human bre
45	43	71.7	2333	12 AD157685	Ad157685 Human bre

ALIGNMENTS

RESULT 1
AAV48320
ID AAV48320 standard; DNA; 60 BP.
XX
XX AAV48320;
XX
XX 20-NOV-1998 (first entry)
XX
XX
DE Nucleotide sequence encoding MUC1 tandem repeat unit R5.
KW ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
KW tumour; tumour-associated antigen.
XX
XX Homo sapiens.
OS
XX
XX
XX WO9837095-A2.
XX
XX 27-AUG-1998.
XX
XX 24-FEB-1998; 98WO-US003693.
XX
XX 24-FEB-1997; 97US-0038253P.
XX
XX (THER-) THERION BIOLOGICS CORP.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX (DAND) DANA-FARBER CANCER INST INC.
XX
XX Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;
XX
XX WPI; 1998-467492/40.
XX
XX New recombinant pox virus for tumour therapy - comprises DNA encoding an
XX immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
XX
XX Disclosure; Page 11; 42pp; English.
XX
XX The MUC1 tandem repeat units AAV48317-V48325 were used to create an
XX immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
XX (RPV). The RPV was used in a pharmaceutical composition also containing
XX an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The
XX recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
XX does not undergo significant genetic deletion, thereby providing an
XX unexpectedly stable and immunogenic pox virus. They can be used to
XX prevent or treat tumours expressing MUC1 tumour-associated antigens

XX Sequence 60 BP; 9 A; 27 C; 15 G; 9 T; 0 U; 0 Other;
SQ

Query Match 100.0%; Score 60; DB 2; Length 60;
Best Local Similarity 100.0%; Pred. No. 1.9e-10;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GGTTGACGCGCCCTCTGCTACCGGTGTACATCCGCCCGATACGAGCCGAGCCCT 60
Db 1 GGTTGACGCGCCCTCTGCTACCGGTGTACATCCGCCCGATACGAGCCGAGCCCT 60

RESULT 2
AAV48316
ID AAV48316 standard; cDNA; 60 BP.

XX AAV48316;

DT 20-NOV-1998 (first entry)

DE Nucleotide sequence encoding MUC1 tandem repeat unit.

XX ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
KM tumour; tumour-associated antigen.

XX Homo sapiens.

XX Key Location/Qualifiers
FT 1..60
FT CDS /tag= a
FT /product= "MUC1 tandem repeat unit"

PN W09837095-A2.

XX 27-AUG-1998.

PF 24-FEB-1998; 98WO-US003693.

PR 24-FEB-1997; 97US-0038253P.

PA (THER-) THERION BIOLOGICS CORP.
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA (DAND) DANA FARBER CANCER INST INC.

PI Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;
WP1; 1998-467492/40.
DR P-PSDB; AAW77229.

DR P-PSDB; AAW77229.

PT New recombinant pox virus for tumour therapy - comprises DNA encoding an
PT immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
XX

PS Example 1; Page 20; 42pp; English.

XX The MUC1 tandem repeat unit was used to create an immunogenic mini-MUC1
CC fragment for inclusion in a recombinant pox virus (RPV). The RPV was used
CC in a pharmaceutical composition also containing an immunomodulator to
CC generate MUC1 specific cytotoxic T-lymphocytes. The recombinant pox virus
CC therefore encodes an immunogenic MUC1 fragment that does not undergo
CC significant genetic deletion, thereby providing an unexpectedly stable
CC and immunogenic pox virus. They can be used to prevent or treat tumours
CC expressing MUC1 tumour-associated antigens
XX

SQ Sequence 60 BP; 7 A; 33 C; 16 G; 4 T; 0 U; 0 Other;

Query Match 71.7%; Score 43; DB 2; Length 60;
Best Local Similarity 83.1%; Pred. No. 7.9e-05;
Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Oy 1 GGTTGACGCGCCCTCTGCTACCGGTGTACATCCGCCCGATACGAGCCGAGCCCT 59
Db 1 GGCTCCACGCGCCCTCTGCTACCGGTGTACATCCGCCCGATACGAGCCGAGCCCT 59

RESULT 3

AA90579/C
ID AA90579 standard; cDNA; 309 BP.

XX AA90579;

DT 27-AUG-2003 (revised)

DT 25-MAR-2003 (revised)

DT 04-DEC-1989 (first entry)

DE PDF9.3 cDNA insert.

XX PDF9.3; human DF3 breast carcinoma-associated antigen epitope.

XX Human MCF-7 breast carcinoma cells; 'ATCC HTB22'.

XX W08907107-A.

XX 10-AUG-1989.

PF 29-JAN-1988; 88US-00149831.

PR 29-JAN-1988; 88US-00149831.

PA (DANA-) DANA-FARBER CANCER.

PI Kufe DW;

DR WP1; 1989-248989/34.

DR P-PSDB; AAP91045; AAP91053; AAP91054; AAP90146.

PT Recombinant polypeptide(s) - contains DF 3 breast carcinoma antigen
PT epitope and useful as assay reagents, and encoding DNA sequences.
XX

PS Claim 1; Fig 4; 31pp; English.

XX The sequence encodes a carbohydrate-free polypeptide contg. DF3 breast
CC carcinoma antigen epitope. Useful as a competitive binding assay reagent
CC and improves diagnosis. The cDNA consists of nearly identical 60 BP
CC tandem repeats which are 85% GC-rich. See also AAP91045, AAP91046,
CC AAP91053, AAP91054. (Updated on 25-MAR-2003 to correct PA field.)
XX (Updated on 27-AUG-2003 to correct OS field.)

SQ Sequence 309 BP; 21 A; 92 C; 160 G; 36 T; 0 U; 0 Other;

Query Match 71.7%; Score 43; DB 1; Length 309;
Best Local Similarity 83.1%; Pred. No. 8.9e-05;
Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Oy 1 GGTTGACGCGCCCTCTGCTACCGGTGTACATCCGCCCGATACGAGCCGAGCCCT 59
Db 266 GGCTCCACGCGCCCTCTGCTACCGGTGTACATCCGCCCGATACGAGCCGAGCCCT 208

RESULT 4

AD157712
ID AD157712 standard; cDNA; 1194 BP.

XX AD157712;

DT 22-APR-2004 (first entry)

XX Human breast specific nucleic acid (BSNA) #83.

XX Human; breast specific nucleic acid; BSNA; gene; ss; metastasis;
KM breast cancer; cytostatic.

XX Homo sapiens.

XX W02003106648-A2.

PD 24-DEC-2003.


```
XX 16-JUN-2003; 2003WO-US018934.
PF 14-JUN-2002; 2002US-0389327P.
XX (DIAD-) DIADEXUS INC.
XX Salceda S, Macina RA, Turner LR, Sun Y, Liu C;
PI WPI; 2004-082185/08.
XX P-PSDB; ADI57782.
DR Novel isolated polypeptide comprising breast specific protein sequences,
PT useful for diagnosing or monitoring presence and metastases of breast
PT cancer in patient.
XX Claim 1; SEQ ID NO 83; 370bp; English.
XX The invention relates to human breast specific nucleic acids (BSNA) and
CC the breast specific proteins (BSP) they encode. The nucleic acids are
CC useful for determining the presence of a BSNA in a sample which involves
CC contacting the sample with a BSNA under conditions in which the BSNA will
CC selectively hybridise to a BSNA in the sample, and detecting the
CC hybridisation. The nucleic acids are useful for determining the presence
CC of a BSP in a sample which involves contacting the sample with suitable
CC reagent under conditions in which the reagent will selectively interact
CC with the BSP, and detecting the interaction of the reagent with a BSP in
CC the sample. The nucleic acids and proteins are useful for diagnosing or
CC monitoring the presence and metastases of breast cancer in a patient,
CC which involves determining an amount of nucleic acid or protein and
CC comparing the determined amount of nucleic acid or protein in the sample
CC of the patient to the amount of a breast specific marker in a normal
CC control, where a difference in the determined amount in the sample
CC compared to the amount in the control is associated with the presence of
CC breast cancer. The sequences are useful for treating a patient with
CC breast cancer, involving administering a composition consisting of a BSNA
CC or a BSP to a patient, where the administration induces an immune
CC response against the breast cancer cell expressing the BSNA or BSP. This
CC sequence represents a human BSNA of the invention.
XX SQ Sequence 1194 BP; 243 A; 438 C; 275 G; 238 T; 0 U; 0 Other;
Query Match 71.7%; Score 43; DB 12; Length 1194;
Best Local Similarity 83.1%; Pred. No. 9.9e-05;
Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 1 GGTTCAGCGGCCCCCTCAGGTGAACATCCGCCGAGTACAGACGGCCCC 59
DB 596 GGCTCACC GCCGCCGCCAGCCACGCTCACCCTCGGCCCGGACACGAGCGGCC 654
RESULT 5
ADIS7693
ID ADIS7693 standard; cDNA; 1378 BP.
XX ADIS7693;
AC 22-APR-2004 (first entry)
XX Human breast specific nucleic acid (BSNA) #64.
DE Human breast specific nucleic acid; BSNA; gene; ss; metastasis;
XX Human; breast specific nucleic acid; BSNA; gene; ss; metastasis;
KM breast cancer; cytosolic.
XX Homo sapiens.
OS WO2003106648-A2.
XX PN 24-DEC-2003.
XX PD 16-JUN-2003; 2003WO-US018934.
XX PF 14-JUN-2002; 2002US-0389327P.
```

```
XX (DIAD-) DIADEXUS INC.
PA Salceda S, Macina RA, Turner LR, Sun Y, Liu C;
XX WPI; 2004-082185/08.
XX P-PSDB; ADI57765.
DR Novel isolated polypeptide comprising breast specific protein sequences,
PT useful for diagnosing or monitoring presence and metastases of breast
PT cancer in patient.
XX Claim 1; SEQ ID NO 64; 370bp; English.
XX The invention relates to human breast specific nucleic acids (BSNA) and
CC the breast specific proteins (BSP) they encode. The nucleic acids are
CC useful for determining the presence of a BSNA in a sample which involves
CC contacting the sample with a BSNA under conditions in which the BSNA will
CC selectively hybridise to a BSNA in the sample, and detecting the
CC hybridisation. The nucleic acids are useful for determining the presence
CC of a BSP in a sample which involves contacting the sample with suitable
CC reagent under conditions in which the reagent will selectively interact
CC with the BSP, and detecting the interaction of the reagent with a BSP in
CC the sample. The nucleic acids and proteins are useful for diagnosing or
CC monitoring the presence and metastases of breast cancer in a patient,
CC which involves determining an amount of nucleic acid or protein and
CC comparing the determined amount of nucleic acid or protein in the sample
CC of the patient to the amount of a breast specific marker in a normal
CC control, where a difference in the determined amount in the sample
CC compared to the amount in the control is associated with the presence of
CC breast cancer. The sequences are useful for treating a patient with
CC breast cancer, involving administering a composition consisting of a BSNA
CC or a BSP to a patient, where the administration induces an immune
CC response against the breast cancer cell expressing the BSNA or BSP. This
CC sequence represents a human BSNA of the invention.
XX SQ Sequence 1378 BP; 295 A; 489 C; 316 G; 278 T; 0 U; 0 Other;
Query Match 71.7%; Score 43; DB 12; Length 1378;
Best Local Similarity 83.1%; Pred. No. 0.0001;
Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 1 GGTTCAGCGGCCCCCTCAGGTGAACATCCGCCGAGTACAGACGGCCCC 59
DB 596 GGCTCACC GCCGCCGCCAGCCACGCTCACCCTCGGCCCGGACACGAGCGGCC 654
RESULT 6
ADO23180/c
ID ADO23180 standard; RNA; 1424 BP.
XX ADO23180;
AC 12-AUG-2004 (first entry)
XX Antisense human MUC1 mucin glycoprotein RNA (coding sequence) Segid 75.
DE human; MUC1; mucin glycoprotein; cancer; chemotherapeutic; MUC1/BCD; ss;
XX antisense.
XX Homo sapiens.
OS WO2004044160-A2.
XX PN 27-MAY-2004.
XX PD 12-NOV-2003; 2003WO-US035848.
XX PF 13-NOV-2002; 2002US-00293391.
XX PR 29-MAY-2003; 2003US-00447839.
XX PA (DAND ) DANA FARBER CANCER INST INC.
XX (ILEX-) ILEX PROD INC.
```

```
XX
PI Kufe DW, Kharbada S, Weitman SD;
XX
DR WPI; 2004-420304/39.
XX
PT Double-stranded RNA complex useful for inhibiting proliferation of cancer
PT cell expressing MUC1 mucin glycoprotein, comprises first and second RNA
XX sequences.
XX
PS Disclosure; SEQ ID NO 75; 112bp; English.
XX
CC This invention relates to novel modulators of the human MUC1 mucin
CC glycoprotein for use in cancer therapeutics, where MUC1 is a protein that
CC acts to inhibit the apoptotic response to genotoxic stress caused by
CC chemotherapeutic agents. In particular, it refers to modulators of the
CC MUC1 extracellular domain (MUC1/ECED). The method refers to using double-
CC stranded RNA complexes as MUC1 interference RNA compositions such that
CC MUC1 expression is inhibited, which in turn inhibits cancer cell
CC proliferation. The present invention describes screening assays to
CC identify compounds that inhibit the binding of various MUC1 ligands such
CC as neutregulin 2, as well as agonists, antagonists and antibodies thereof.
CC Furthermore, it provides MUC1 antisense, small interfering RNA (siRNA)
CC and small molecules in combination with chemotherapeutic agents that are
CC useful in the field of cancer therapy. This polynucleotide sequence is
CC the antisense human MUC1 RNA of the invention.
XX
SQ Sequence 1424 BP; 309 A; 325 C; 495 G; 0 T; 295 U; 0 Other;
XX
Query Match 71.7%; Score 43; DB 12; Length 1424;
Best Local Similarity 83.1%; Pred. No. 0.0001;
Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
XX
QY 1 GGTTCGACGGCGCCCTGCTCAAGGTAAACATCCCGGATGCCAGACCGGGCCCC 59
Db 1041 GGCTCCACCGCGCCCGCCAGCCAGGTGTCACCTCGGCCCCGACACAGCGCGCCCC 983
XX
RESULT 7
ABL60159
ID ABL60159 standard; cDNA; 1428 BP.
XX
AC ABL60159;
XX
DT 22-JUL-2002 (first entry)
XX
DE Human MUC1 encoding cDNA SEQ ID NO 2.
XX
KW Human; mucin 1; MUC1; transmembrane protein; SNP; cancer; cytostatic;
KW single nucleotide polymorphism; haplotyping; genotyping; drug;
XX antiinflammatory; gene; ss.
XX
OS Homo sapiens.
XX
XX
XX Key Location/Qualifiers
XX CDS 1..1428
XX FT /*tag= a
XX FT /product= "MUC1"
XX FT replace(1009,A)
XX FT /*tag= b
XX FT /standard_name= "Single nucleotide polymorphism"
XX FT /note= "SNP allelic variation results in Val substituted
XX by Met at position 337 of the MUC1 protein (AB077476)"
XX
XX MO200226765-A2.
XX
XX 04-APR-2002.
XX
XX 25-SEP-2001; 2001WO-US030151.
XX
XX 28-SEP-2000; 2000US-0236113P.
XX
XX (GENA-) GENAISSANCE PHARM INC.
```

```
PI Chew A, Koshy B;
XX
DR WPI; 2002-405042/43.
XX
DR P-PSDB; ABB77476.
XX
XX
PT New genetic variants of mucin 1, Transmembrane gene, useful in studying
PT expression and function of protein encoded by the gene and for screening
XX drugs to treat diseases e.g. cancer.
XX
XX Claim 23; Fig 2; 75pp; English.
XX
XX
CC The invention relates to a polynucleotide (ABL60158, ABL60159) encoding
CC mucin 1/MUC1 (ABB77476), Transmembrane isogene. The invention describes
CC novel genetic variants of the MUC1 gene. The invention is useful for
CC haplotyping/genotyping the MUC1 gene in an individual and identifying an
CC association between a trait and at least one of the haplotypes or
CC haplotype pairs of MUC1 gene. MUC1 is useful for studying the expression
CC and function of MUC1 and expressing MUC1 protein for use in screening for
CC candidate drugs to treat diseases related to MUC1 activity and in
CC studying the effect of the variation on the biological activity of MUC1
CC as well as on the binding affinity of candidate drugs targeting MUC1 for
CC the treatment of e.g. cancer. MUC1 is further used by the pharmaceutical
CC research scientist to validate MUC1 as a candidate target for and in
CC design of clinical trials of candidate drugs for, treating a specific
CC condition drugs or disease predicted to be associated with MUC1 activity.
CC MUC1 antibodies are useful in a variety of diagnostic and prognostic
XX formats and therapeutic methods
XX
SQ Sequence 1428 BP; 296 A; 496 C; 327 G; 309 T; 0 U; 0 Other;
XX
Query Match 71.7%; Score 43; DB 6; Length 1428;
Best Local Similarity 83.1%; Pred. No. 0.0001;
Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
XX
QY 1 GGTTCGACGGCGCCCTGCTCAAGGTAAACATCCCGGATGCCAGACCGGGCCCC 59
Db 385 GGCTCCACCGCGCCCGCCAGCCAGGTGTCACCTCGGCCCCGACACAGCGCGCCCC 443
XX
RESULT 8
ADO23125
ID ADO23125 standard; RNA; 1428 BP.
XX
AC ADO23125;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human MUC1 mucin glycoprotein RNA (coding sequence) SegID 20.
XX
KW human; MUC1; mucin glycoprotein; cancer; chemotherapeutic; MUC1/ECED; ss.
XX
XX Homo sapiens.
XX
XX
XX WO2004044160-A2.
XX
XX 27-MAY-2004.
XX
XX 12-NOV-2003; 2003WO-US035848.
XX
XX 13-NOV-2002; 2002US-00293391.
XX
XX 29-MAY-2003; 2003US-00447839.
XX
XX (DAND ) DANA FARBER CANCER INST INC.
XX
XX (ILEX-) ILEX PROD INC.
XX
XX Kufe DW, Kharbada S, Weitman SD;
XX
XX WPI; 2004-420304/39.
XX
XX Double-stranded RNA complex useful for inhibiting proliferation of cancer
XX PT cell expressing MUC1 mucin glycoprotein, comprises first and second RNA
XX PT sequences.
```

PS Claim 2; SEQ ID NO 20; 112pp; English.

XX This invention relates to novel modulators of the human MUC1 mucin glycoprotein for use in cancer therapeutics, where MUC1 is a protein that acts to inhibit the apoptotic response to genotoxic stress caused by chemotherapeutic agents. In particular, it refers to modulators of the MUC1 extracellular domain (MUC1/ECF). The method refers to using double-stranded RNA complexes as MUC1 interference RNA compositions such that MUC1 expression is inhibited, which in turn inhibits cancer cell proliferation. The present invention describes screening assays to identify compounds that inhibit the binding of various MUC1 ligands such as neueregulin 2, as well as agonists, antagonists and antibodies thereof. Furthermore, it provides MUC1 antisense, small interfering RNA (siRNA) and small molecules in combination with chemotherapeutic agents that are useful in the field of cancer therapy. This polynucleotide sequence is the human MUC1 RNA of the invention.

CC Sequence 1428 BP; 296 A; 496 C; 327 G; 0 T; 309 U; 0 Other;

XX Query Match 71.7%; Score 43; DB 12; Length 1428;

XX Best Local Similarity 76.3%; Pred. No. 0.0001;

XX Matches 45; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

OY 1 GGTTCGACGGCCCCCTGCTCAAGGTATACATCCGCCGATACGAGCCGCC 59

DB 385 GGCTCACCAGCCCCCCCCAGCCAGGUGUACUCGCCCCGACACGAGCCGCC 443

RESULT 9

ADP32627

ID ADF32627 standard; DNA; 1457 BP.

XX ADF32627;

XX 26-FEB-2004 (first entry)

XX Plasmid JMW358 MUC-1 nucleotide sequence.

XX MUC-1 antigen; immune response; MUC-1; variable number of tandem repeat;

XX VNR; repeat unit; tumour; metastasis; cytostatic; vaccine; gene therapy;

XX gene; da.

XX Synthetic.

XX WO2003100660-A2.

XX 04-DEC-2003.

XX 23-MAY-2003; 2003WO-EP005594.

XX 24-MAY-2002; 2002GB-00012046.

XX (GLAXO) GLAXO GROUP LTD.

XX Burden N, Ellis JH, Hamblin PA;

XX WPI; 2004-042811/04.

XX New nucleic acid molecule encoding a MUC-1 antigen, useful for preparing a composition for treating or preventing tumors or metastases.

XX Example; Fig 3; 66pp; English.

XX The present invention describes a nucleic acid molecule which encodes a MUC-1 antigen. The nucleic acid is capable of raising an immune response in vivo, has reduced susceptibility to recombination than full-length MUC-1 and comprises between 1 and 15 variable number of tandem repeats (VNR) perfect repeat units. Also described: (1) a plasmid comprising the DNA molecule; (2) a protein encoded by the nucleic acid; (3) a pharmaceutical composition comprising the nucleic acid, plasmid or protein and an excipient, diluent or carrier; and (4) a method of treating or preventing tumours or metastases. A MUC1 antigen has cytostatic activity, and can be used in vaccines, and in gene therapy.

CC The nucleic acid is useful for preparing a composition for treating or preventing tumours or metastases. The present sequence is used in the exemplification of the present invention.

XX Sequence 1457 BP; 302 A; 506 C; 334 G; 315 T; 0 U; 0 Other;

XX Query Match 71.7%; Score 43; DB 12; Length 1457;

XX Best Local Similarity 83.1%; Pred. No. 0.0001;

XX Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 1 GGTTCGACGGCCCCCTGCTCAAGGTATACATCCGCCGATACGAGCCGCC 59

DB 403 GGCTCACCAGCCCCCCCCAGCCAGGUGUACUCGCCCCGACACGAGCCGCC 461

RESULT 10

AAV48329

ID AAV48329 standard; cDNA; 1527 BP.

XX AAV48329;

XX 20-NOV-1998 (first entry)

XX MiniMUC1 Gene.

XX ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;

XX tumour; tumour-associated antigen.

XX Homo sapiens.

XX Key Location/Qualifiers

XX FT 1..1527 /tag= a

XX FT CDS /product= "MiniMUC1 protein"

XX WO9837095-A2.

XX 27-AUG-1998.

XX 24-FEB-1998; 98WO-US003693.

XX 24-FEB-1997; 97US-0038253P.

XX (THER-) THERION BIOLOGICS CORP.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX (DAND) DANA FARBER CANCER INST INC.

XX Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;

XX WPI; 1998-467492/40.

XX P-PSDB; AAW77233.

XX New recombinant pox virus for tumour therapy - comprises DNA encoding an immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.

XX Example 1; Page 21-22; 42pp; English.

XX The immunogenic mini-MUC1 fragment was created from MUC1 tandem repeat units for inclusion in a recombinant pox virus (RPV). The RPV was used in a pharmaceutical composition also containing an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The recombinant pox virus therefore encodes an immunogenic MUC1 fragment that does not undergo significant genetic deletion, thereby providing an unexpectedly stable and immunogenic pox virus. They can be used to prevent or treat tumours expressing MUC1 tumour-associated antigens

XX Sequence 1527 BP; 296 A; 573 C; 351 G; 307 T; 0 U; 0 Other;

XX Query Match 71.7%; Score 43; DB 2; Length 1527;

XX Best Local Similarity 83.1%; Pred. No. 0.0001;

XX Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 1 GGTTCGACGGCCCCCTGCTCAAGGTATACATCCGCCGATACGAGCCGCC 59

Db 226 GGCTCCACCGCCCCCGACCGGTGTACCTCGCCCCGACACCGCGGCCCC 284

RESULT 11
ADK70370
ID ADK70370 standard; cDNA; 1614 BP.

ADK70370;

06-MAY-2004 (first entry)

Respiratory disease differentially expressed cDNA #106.

de; gene; cytostatic; respiratory; antiasthmatic; gene therapy;
KW differential gene expression; respiratory disorder; lung cancer;
XX chronic obstructive pulmonary disease; emphysema; asthma.

OS Homo sapiens.

PN WO2003101283-A2.

PD 11-DEC-2003.

PF 02-JUN-2003; 2003WO-US017409.

PR 04-JUN-2002; 2002US-0386005P.

PA (INCY-) INCYTE CORP.

PI Rickert PK, Krasnow R;

DR WPI; 2004-042945/04.

XX New combination comprising cDNAs and proteins that are differentially
XX expressed in respiratory disorders, useful for diagnosing or treating
XX respiratory diseases e.g. lung cancer, chronic obstructive pulmonary
XX diseases or asthma.

PS Claim 1; SEQ ID NO 106; 343bp; English.

XX The invention relates to cDNA sequences that are differentially expressed
XX in respiratory disorders or their complements or encoded proteins. The
XX cDNAs and proteins are useful for diagnosing, treating or monitoring
XX treatment of a subject with a respiratory disease including lung cancer,
XX chronic obstructive pulmonary diseases, emphysema or asthma. The protein
XX is also useful for screening molecules or compounds to identify at least
XX one ligand which specifically binds the protein. It is also useful for
XX preparing and purifying a polyclonal or monoclonal antibody. This
XX sequence corresponds to a cDNA of the invention.

SQ Sequence 1614 BP; 329 A; 556 C; 374 G; 353 T; 0 U; 2 Other;

Query Match 71.7%; Score 43; DB 12; Length 1614;
Best Local Similarity 83.1%; Pred. No. 0.0001;

Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 GGTTCAGCGCCCCCTGCTCAGGTGAACATCGCCCCGGATACCAAGCCGCC 59
Db 434 GGCTCCACCGCCCCCGACCGGTGTACCTCGCCCCGACACCGCGGCCCC 492

RESULT 12
AD157708
ID AD157708 standard; cDNA; 1630 BP.

AD157708;

22-APR-2004 (first entry)

Human breast specific nucleic acid (BSNA) #79.

Human; breast specific nucleic acid; BSNA; gene; ss; metastasis;

KW breast cancer; cytostatic.

OS Homo sapiens.

PN WO2003106648-A2.

PD 24-DEC-2003.

PF 16-JUN-2003; 2003WO-US018934.

PR 14-JUN-2002; 2002US-0389327P.

PA (DIAD-) DIADEXUS INC.

PI Salceda S, Macina RA, Turner LR, Sun Y, Liu C;

DR WPI; 2004-082185/08.

DR P-PSDB; AD157779.

PT Novel isolated polypeptide comprising breast specific protein sequences,
PT useful for diagnosing or monitoring presence and metastases of breast
PT cancer in patient.

PS Claim 1; SEQ ID NO 79; 370bp; English.

XX The invention relates to human breast specific nucleic acids (BSNA) and
XX the breast specific proteins (BSP) they encode. The nucleic acids are
XX useful for determining the presence of a BSNA in a sample which involves
XX contacting the sample with a BSNA under conditions in which the BSNA will
XX selectively hybridise to a BSNA in the sample, and detecting the
XX hybridisation. The nucleic acids are useful for determining the presence
XX of a BSP in a sample which involves contacting the sample with suitable
XX reagent under conditions in which the reagent will selectively interact
XX with the BSP, and detecting the interaction of the reagent with a BSP in
XX the sample. The nucleic acids and proteins are useful for diagnosing or
XX monitoring the presence and metastases of breast cancer in a patient,
XX which involves determining an amount of nucleic acid or protein and
XX comparing the determined amount of nucleic acid or protein in the sample
XX of the patient to the amount of a breast specific marker in a normal
XX control, where a difference in the determined amount in the sample
XX compared to the amount in the control is associated with the presence of
XX breast cancer. The sequences are useful for treating a patient with
XX breast cancer, involving administering a composition consisting of a BSNA
XX or a BSP to a patient, where the administration induces an immune
XX response against the breast cancer cell expressing the BSNA or BSP. This
XX sequence represents a human BSNA of the invention.

SQ Sequence 1630 BP; 322 A; 584 C; 405 G; 319 T; 0 U; 0 Other;

Query Match 71.7%; Score 43; DB 12; Length 1630;
Best Local Similarity 83.1%; Pred. No. 0.0001;

Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 GGTTCAGCGCCCCCTGCTCAGGTGAACATCGCCCCGGATACCAAGCCGCC 59
Db 536 GGCTCCACCGCCCCCGACCGGTGTACCTCGCCCCGACACCGCGGCCCC 654

RESULT 13
AD157689
ID AD157689 standard; cDNA; 1634 BP.

AD157689;

22-APR-2004 (first entry)

Human breast specific nucleic acid (BSNA) #60.

Human; breast specific nucleic acid; BSNA; gene; ss; metastasis;
breast cancer; cytostatic.

Homo sapiens.

PN WO2003106648-A2.
XX
PD 24-DEC-2003.
XX
XX 16-JUN-2003; 2003WO-US018934.
PF
XX 14-JUN-2002; 2002US-0389327P.
PR
XX (DIAD-) DIADEXUS INC.
PA
PI Salceda S, Macina RA, Turner LR, Sun Y, Liu C;
XX
XX MPI; 2004-082185/08.
DR
P-PSDB; AD157761.
PT
PT Novel isolated polypeptide comprising breast specific protein sequences,
PT useful for diagnosing or monitoring presence and metastases of breast
PT cancer in patient.
XX
PS Claim 1; SEQ ID NO 60; 370pp; English.
XX
XX The invention relates to human breast specific nucleic acids (BSNA) and
CC the breast specific proteins (BSP) they encode. The nucleic acids are
CC useful for determining the presence of a BSNA in a sample which involves
CC contacting the sample with a BSNA under conditions in which the BSNA will
CC selectively hybridise to a BSNA in the sample, and detecting the presence
CC of a BSP in a sample which involves contacting the sample with suitable
CC reagent under conditions in which the reagent will selectively interact
CC with the BSP, and detecting the interaction of the reagent with a BSP in
CC the sample. The nucleic acids and proteins are useful for diagnosing or
CC monitoring the presence and metastases of breast cancer in a patient,
CC which involves determining an amount of nucleic acid or protein and
CC comparing the determined amount of nucleic acid or protein in the sample
CC of the patient to the amount of a breast specific marker in a normal
CC control, where a difference in the determined amount in the sample
CC compared to the amount in the control is associated with the presence of
CC breast cancer. The sequences are useful for treating a patient with
CC breast cancer, involving administering a composition consisting of a BSNA
CC or a BSP to a patient, where the administration induces an immune
CC response against the breast cancer cell expressing the BSNA or BSP. This
CC sequence represents a human BSNA of the invention.
XX
SQ Sequence 1634 BP; 349 A; 541 C; 405 G; 337 T; 0 U; 2 Other;
XX
XX Query Match 71.7%; Score 43; DB 12; Length 1634;
Best Local Similarity 83.1%; Pred. No. 0.0001;
Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 1 GGTTCGACGGCCCCCTGCTCAGCGTGTACATCCGCCGAGATCCAGACGGCCCC 59
DB 596 GGCTCCACCGCCCCCCAGCCAGCGGTGTACCTCGGCCCGGACACGAGCGGCCCC 654
RESULT 14
AD157686
ID AD157686 standard; cDNA; 1712 BP.
XX
XX AD157686;
XX
XX 22-APR-2004 (first entry)
XX
XX Human breast specific nucleic acid (BSNA) #57.
DE
XX Human breast specific nucleic acid (BSNA) #57.
XX
XX Human; breast specific nucleic acid; BSNA; gene; ss; metastasis;
KM breast cancer; cytosolic.
XX
XX Homo sapiens.
OS
XX WO2003106648-A2.
XX
XX 24-DEC-2003.
XX
XX

PF 16-JUN-2003; 2003WO-US018934.
XX
XX 14-JUN-2002; 2002US-0389327P.
XX
XX (DIAD-) DIADEXUS INC.
PA
PI Salceda S, Macina RA, Turner LR, Sun Y, Liu C;
XX
XX MPI; 2004-082185/08.
DR
P-PSDB; AD157758.
PT
PT Novel isolated polypeptide comprising breast specific protein sequences,
PT useful for diagnosing or monitoring presence and metastases of breast
PT cancer in patient.
XX
PS Claim 1; SEQ ID NO 57; 370pp; English.
XX
XX The invention relates to human breast specific nucleic acids (BSNA) and
CC the breast specific proteins (BSP) they encode. The nucleic acids are
CC useful for determining the presence of a BSNA in a sample which involves
CC contacting the sample with a BSNA under conditions in which the BSNA will
CC selectively hybridise to a BSNA in the sample, and detecting the presence
CC of a BSP in a sample which involves contacting the sample with suitable
CC reagent under conditions in which the reagent will selectively interact
CC with the BSP, and detecting the interaction of the reagent with a BSP in
CC the sample. The nucleic acids and proteins are useful for diagnosing or
CC monitoring the presence and metastases of breast cancer in a patient,
CC which involves determining an amount of nucleic acid or protein and
CC comparing the determined amount of nucleic acid or protein in the sample
CC of the patient to the amount of a breast specific marker in a normal
CC control, where a difference in the determined amount in the sample
CC compared to the amount in the control is associated with the presence of
CC breast cancer. The sequences are useful for treating a patient with
CC breast cancer, involving administering a composition consisting of a BSNA
CC or a BSP to a patient, where the administration induces an immune
CC response against the breast cancer cell expressing the BSNA or BSP. This
CC sequence represents a human BSNA of the invention.
XX
SQ Sequence 1712 BP; 327 A; 604 C; 435 G; 346 T; 0 U; 0 Other;
XX
XX Query Match 71.7%; Score 43; DB 12; Length 1712;
Best Local Similarity 83.1%; Pred. No. 0.0001;
Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 1 GGTTCGACGGCCCCCTGCTCAGCGTGTACATCCGCCGAGATCCAGACGGCCCC 59
DB 596 GGCTCCACCGCCCCCCAGCCAGCGGTGTACCTCGGCCCGGACACGAGCGGCCCC 654
RESULT 15
AD157669
ID AD157669 standard; cDNA; 1738 BP.
XX
XX AD157669;
XX
XX 22-APR-2004 (first entry)
XX
XX Human breast specific nucleic acid (BSNA) #40.
DE
XX Human breast specific nucleic acid (BSNA) #40.
XX
XX Human; breast specific nucleic acid; BSNA; gene; ss; metastasis;
KM breast cancer; cytosolic.
XX
XX Homo sapiens.
OS
XX WO2003106648-A2.
XX
XX 24-DEC-2003.
XX
XX 16-JUN-2003; 2003WO-US018934.
XX
XX 14-JUN-2002; 2002US-0389327P.
XX
XX

PA (DIAD-) DIADEXUS INC.
XX
PI Salceda S, Macina RA, Turner LR, Sun Y, Liu C;
XX WPI; 2004-082185/08.
DR P-PSDB; ADI57743.
XX
XX Novel isolated polypeptide comprising breast specific protein sequences,
PT useful for diagnosing or monitoring presence and metastases of breast
PT cancer in patient.
XX
XX Claim 1; SEQ ID NO 40; 370pp; English.
XX
XX The invention relates to human breast specific nucleic acids (BSNA) and
CC the breast specific proteins (BSP) they encode. The nucleic acids are
CC useful for determining the presence of a BSNA in a sample which involves
CC contacting the sample with a BSNA under conditions in which the BSNA will
CC selectively hybridise to a BSNA in the sample, and detecting the
CC hybridisation. The nucleic acids are useful for determining the presence
CC of a BSP in a sample which involves contacting the sample with suitable
CC reagent under conditions in which the reagent will selectively interact
CC with the BSP, and detecting the interaction of the reagent with a BSP in
CC the sample. The nucleic acids and proteins are useful for diagnosing or
CC monitoring the presence and metastases of breast cancer in a patient,
CC which involves determining an amount of nucleic acid or protein and
CC comparing the determined amount of nucleic acid or protein in the sample
CC of the patient to the amount of a breast specific marker in a normal
CC control, where a difference in the determined amount in the sample
CC compared to the amount in the control is associated with the presence of
CC breast cancer. The sequences are useful for treating a patient with
CC breast cancer, involving administering a composition consisting of a BSNA
CC or a BSP to a patient, where the administration induces an immune
CC response against the breast cancer cell expressing the BSNA or BSP. This
CC sequence represents a human BSNA of the invention.
XX
SQ Sequence 1738 BP; 342 A; 642 C; 399 G; 355 T; 0 U; 0 Other;

Query Match 71.7%; Score 43; DB 12; Length 1738;
Best Local Similarity 83.1%; Pred. No. 0.0001;
Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 GGTTCGACGGCCCCCGCTGCTACCGGTGTAATCCGCGGATACGAGACCGGCCCC 59
|||
Db 458 GGCTCCACCGCCCCCGAGCCAGGTGTCCTCGGCCGACACGAGCGGCGCCCC 516
|||

Search completed: January 15, 2005, 20:36:09
Job time : 170.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2005, 17:40:52 ; Search time 722.8 Seconds
(without alignments)
3925.542 Million cell updates/sec

Title: US-10-057-136-7

Perfect score: 60

Sequence: 1 ggttcgacgagcccccctgc.....cggatcacgagccgcct 60

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 2364849745 residues

9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Genemb1:*
1: gb_ba:*
2: gb_hlg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pac:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sbs:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	2297	6	BD272907 A recombi
2	43	71.7	120	6	AX192396 Sequence
3	43	71.7	120	6	BD000571 Human pol
4	43	71.7	1414	12	AF423031 Synthetic
5	43	71.7	1455	6	CO715242 Sequence
6	43	71.7	1457	6	AX959914 Sequence
7	43	71.7	1774	6	AX959912 Sequence
8	43	71.7	1774	6	AX959912 Sequence
9	43	71.7	1804	6	AR492318 Sequence
10	43	71.7	1804	6	AX335367 Sequence
11	43	71.7	1804	9	HUMMUCAB
12	43	71.7	1834	12	AF423030
13	43	71.7	1835	6	AX959918 Sequence
14	43	71.7	2135	6	AX959916 Sequence
15	43	71.7	4139	6	CO834017 Sequence
16	43	71.7	4139	6	AX334899 Sequence
17	43	71.7	4139	6	AX335372 Sequence
18	43	71.7	4139	6	AX336712 Sequence
19	43	71.7	4139	6	AX409474 Sequence

20	43	71.7	4139	6	AX440481	AX440481 Sequence
21	43	71.7	4139	9	HUMPANMU	J05582 Human pancr
22	43	71.7	7188	9	AY463543	AY463543 Homo sapi
23	43	71.7	8181	6	AX406624	AX406624 Sequence
24	43	71.7	8181	6	HUMPEM	W61170 Human polyom
25	43	71.7	8186	6	AR492320	AR492320 Sequence
26	43	71.7	133525	9	AL713999	AL713999 Human DNA
27	41.4	69.0	120	6	AX192397	AX192397 Sequence
28	41.4	69.0	120	6	BD000572	BD000572 Human pol
29	41.4	69.0	180	6	E08763	E08763 CDNA encodi
30	41.4	69.0	572	6	AR492319	AR492319 Sequence
31	41.4	69.0	572	9	HUMDF3A	M31823 Human Breaa
32	41.4	69.0	1572	6	AX093798	AX093798 Sequence
33	41.4	69.0	1721	6	CO771290	CO771290 Sequence
34	41.4	69.0	1721	6	AR492306	AR492306 Sequence
35	41.4	69.0	1721	6	AX335860	AX335860 Sequence
36	41.4	69.0	1721	6	AX440427	AX440427 Sequence
37	41.4	69.0	1721	6	AX587588	AX587588 Sequence
38	41.4	69.0	1721	9	HSTEYMA	X52229 Human mRNA
39	41.4	69.0	1971	6	AX963157	AX963157 Sequence
40	41.4	69.0	2037	6	AX963159	AX963159 Sequence
41	41.4	69.0	2238	9	HSEBTA	X52228 Human mRNA
42	41	68.3	491	9	HUMEPISIB1	M32739 Human epist
43	41	68.3	518	6	AR492402	AR492402 Sequence
44	41	68.3	518	9	HUMEPISIA1	M32738 Human epist
45	39.8	66.3	3343	6	AR492403	AR492403 Sequence

ALIGNMENTS

RESULT 1	BD272907	2297 bp	DNA	linear	PAT 17-JUL-2003
LOCUS	BD272907				
DEFINITION	A recombinant vector expressing multiple constitutively molecules and uses thereof.				
ACCESSION	BD272907	1	GI:33082675		
VERSION	BD272907				
KEYWORDS	JP 2002531133-A/1.				
SOURCE	JP 2002531133-A/1.				
ORGANISM	synthetic construct				
REFERENCE	1 (bases 1 to 2297)				
AUTHORS	Schlom,J., Hodge,J. and Panicali,D.				
TITLE	A recombinant vector expressing multiple constitutively molecules and uses thereof				
JOURNAL	Patent: JP 2002531133-A 1 24-SEP-2002;				
COMMENT	THE UNITED STATES OF AMERICA, THERION BIOLOGICS CORP				
OS	Artificial Sequence				
PN	JP 2002531133-A/1				
PD	24-SEP-2002				
PF	12-NOV-1999 JP 2000586927				
PR	09-DEC-1998 US 60/111582				
PI	JEFFREY SCHLOM, JAMES HODGE, DENNIS PANICALI				
PC	CL12N15/02, A61K35/12, A61K35/74, A61K35/76, A61K38/00, A61K39/00, A61K39/12, A61K39/21, A61K39/235, A61K39/245, A61K39/275, A61K39/29, A61K48/00,				
PC	A61K39/12,				
PC	A61K39/125, A61K39/21, A61K39/235, A61K39/245, A61K39/275, A61K39/29, A61K48/00,				
PC	A61P1/04, A61P31/04, A61P31/10, A61P31/12, A61P35/00, A61P37/02, PC A61P37/06,				
PC	CL12N1/15, CL12N1/19, CL12N1/21, CL12N5/10, CL12N7/00, CL12Q1/02, G01N33/53, CL12N15/00,				
PC	CL12N5/00, A61K37/02				
CC	Description of Artificial Sequence: VECTOR SEQUENCE FH				
FT	Location/Qualifiers				
FT	source				
FEATURES	Location/Qualifiers				
source	1..2297				
	/organism="Artificial Sequence".				
	/mol_type="synthetic construct"				
	/db_xref="taxon:32630"				
ORIGIN					

ORIGIN
Query Match
Best Local Similarity 83.1%; Pred. No. 0.0082;
Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

1 GGTTCAGCGCCCCCTGCTCAGCGGTGAACATCCGCCGATACAGACCGGCC 59
|||
385 GGCTCCACCGCCCCCGACCGGTCGACCTCGGCCCGGACACGACCGGCC 443

RESULT 5
LOCUS CQ715242 1455 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 1176 from Patent WO02068579.
ACCESSION CQ715242
VERSION CQ715242.1 GI:42276099
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE
AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
TITLE Kite, such as nucleic acid arrays, comprising a majority of
humanecons or transcripts, for detecting expression and other uses
thereof
Patent: WO 02068579-A 1176 06-SEP-2002;
PE Corporation (NY) (US)
LOCATION/Qualifiers
1..1455
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match
Best Local Similarity 83.1%; Pred. No. 0.0082;
Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

1 GGTTCAGCGCCCCCTGCTCAGCGGTGAACATCCGCCGATACAGACCGGCC 59
|||
412 GGCTCCACCGCCCCCGACCGGTCGACCTCGGCCCGGACACGACCGGCC 470

RESULT 6
LOCUS AX959914 1457 bp DNA linear PAT 14-JAN-2004
DEFINITION Sequence 19 from Patent WO03100060.
ACCESSION AX959914
VERSION AX959914.1 GI:40880143
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE
AUTHORS Burden, N.G., Ellis, J.H. and Hamblin, P.A.
TITLE Muc-1 antigen with reduced number of vntc repeat units
JOURNAL Patent: WO 03100060-A 19 04-DEC-2003;
GLAXO GROUP LIMITED (GB)
LOCATION/Qualifiers
1..1457
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match
Best Local Similarity 83.1%; Pred. No. 0.0082;
Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

1 GGTTCAGCGCCCCCTGCTCAGCGGTGAACATCCGCCGATACAGACCGGCC 59
|||
462 GGCTCCACCGCCCCCGACCGGTCGACCTCGGCCCGGACACGACCGGCC 520

RESULT 9
AR492318

Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

1 GGTTCAGCGCCCCCTGCTCAGCGGTGAACATCCGCCGATACAGACCGGCC 59
|||
403 GGCTCCACCGCCCCCGACCGGTCGACCTCGGCCCGGACACGACCGGCC 461

RESULT 7
LOCUS AX959684 1774 bp DNA linear PAT 14-JAN-2004
DEFINITION Sequence 10 from Patent WO0309193.
ACCESSION AX959684
VERSION AX959684.1 GI:40880030
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE
AUTHORS Burden, N. and Hamblin, P.
TITLE Vaccines
JOURNAL Patent: WO 0309193-A 10 04-DEC-2003;
GLAXO GROUP LIMITED (GB)
LOCATION/Qualifiers
1..1774
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match
Best Local Similarity 83.1%; Pred. No. 0.0078;
Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

1 GGTTCAGCGCCCCCTGCTCAGCGGTGAACATCCGCCGATACAGACCGGCC 59
|||
462 GGCTCCACCGCCCCCGACCGGTCGACCTCGGCCCGGACACGACCGGCC 520

RESULT 8
LOCUS AX959912 1774 bp DNA linear PAT 14-JAN-2004
DEFINITION Sequence 17 from Patent WO03100060.
ACCESSION AX959912
VERSION AX959912.1 GI:40880142
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE
AUTHORS Burden, N.G., Ellis, J.H. and Hamblin, P.A.
TITLE Muc-1 antigen with reduced number of vntc repeat units
JOURNAL Patent: WO 03100060-A 17 04-DEC-2003;
GLAXO GROUP LIMITED (GB)
LOCATION/Qualifiers
1..1774
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match
Best Local Similarity 83.1%; Pred. No. 0.0078;
Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

1 GGTTCAGCGCCCCCTGCTCAGCGGTGAACATCCGCCGATACAGACCGGCC 59
|||
462 GGCTCCACCGCCCCCGACCGGTCGACCTCGGCCCGGACACGACCGGCC 520

RESULT 9
AR492318

```

LOCUS       AR492318             1804 bp    DNA             linear      PAT 15-MAY-2004
DEFINITION   Sequence 17 from patent US 6716627.
ACCESSION    AR492318
VERSION      AR492318.1   GI:47260892
KEYWORDS     .
SOURCE       Unknown.
ORGANISM     Unknown.
REFERENCE    1 (bases 1 to 1804)
AUTHORS      Dobie,K.W.
TITLE        Antisense modulation of mucin 1, transmembrane expression
JOURNAL      Patent: US 6716627-A 17 06-APR-2004;
FEATURES     location/Qualifiers
             source          1..1804
                        /organism="unknown"
                        /mol_type="genomic DNA"

ORIGIN
Query Match      71.7%; Score 43; DB 6; Length 1804;
Best Local Similarity 83.1%; Pred. No. 0.0078;
Matches          49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Cy              1 GGTTCGACGCGCCCGCTGCTACGCTGTAACATCCGCCCGGATACGACGCGCGCC 59
Db              457 GGCTTCACCGCGCCCGCCGACGCTGTCACCTCGGCCCGGACACGACGCGCGCC 515

RESULT 10
LOCUS       AX335367             1804 bp    DNA             linear      PAT 09-JAN-2002
DEFINITION   Sequence 5876 from Patent W00194629.
ACCESSION    AX335367
VERSION      AX335367.1   GI:18126086
KEYWORDS     .
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
REFERENCE    1
AUTHORS      Eukaryotic, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
TITLE        Young,P.E., Augustus,W., Carter,K.C., Ehner,R., Endress,G.,
             Horrigan,S., Soppet,D.R. and Weaver,Z.
             Cancer gene determination and therapeutic screening using signature
             gene sets
JOURNAL      Patent: WO 0194629-A 5876 13-DEC-2001;
FEATURES     location/Qualifiers
             source          1..1804
                        /organism="Homo sapiens"
                        /mol_type="unassigned DNA"
                        /db_xref="taxon:9606"

ORIGIN
Query Match      71.7%; Score 43; DB 6; Length 1804;
Best Local Similarity 83.1%; Pred. No. 0.0078;
Matches          49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Cy              1 GGTTCGACGCGCCCGCTGCTACGCTGTAACATCCGCCCGGATACGACGCGCGCC 59
Db              457 GGCTTCACCGCGCCCGCCGACGCTGTCACCTCGGCCCGGACACGACGCGCGCC 515

RESULT 11
LOCUS       HUMMUCAB             1804 bp    mRNA             linear      PRI 07-JAN-1995
DEFINITION   Human polymorphic epithelial mucin (PEM) mRNA, complete cds.
ACCESSION    J05581.1   GI:188869
KEYWORDS     polymorphic epithelial mucin.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
REFERENCE    1
AUTHORS      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
TITLE        Mammalia, Eutheria, Primates, Catarrhini, Homiidae; Homo.
REFERENCE    1 (bases 1 to 1804)

```

```

AUTHORS      Gendler,S.J., Lancaster,C.A., Taylor-Papadimitriou,J., Duhig,T.,
             Peat,N., Birchell,J., Pemberton,L., Lalani,E.N. and Wilson,D.
TITLE        Molecular cloning and expression of human tumor-associated
             polymorphic epithelial mucin
JOURNAL      J. Biol. Chem. 265 (25), 15286-15293 (1990)
MEDLINE      90368715
PUBMED       1697589
COMMENT      Original source text: Homo sapiens adult adenocarcinoma CDNA to
             mRNA.
             Draft entry and computer-readable sequence for [J. Biol. Chem.
             (1990) In press] kindly submitted
             by S.J. Gendler, 26-JUN-1990.

FEATURES     location/Qualifiers
             source          1..1804
                        /organism="Homo sapiens"
                        /mol_type="mRNA"
                        /db_xref="taxon:9606"
                        /map="Xq26.3-q27.1"
                        /cell_line="BT20"
                        /tissue_type="adenocarcinoma"
                        /dev_stage="adult"
             1..1804
             /gene="MCNAA"
             <1..1804
             /gene="MCNAA"
             /product="PEM mRNA"
             73..1500
             /gene="MCNAA"
             /codon_start=1
             /product="polymorphic epithelial mucin"
             /protein_id="AA59876.1"
             /db_xref="GI:188870"
             /translation="MTPTGQSPFLLILTLVTVNGSHASTPGEKETSATORSS
             VPSSTENAVSMTSVLSHSPGSGSTQGGDTLAAPTEPASGATWGDVTSV
             VTRPAIGSTTPPAHDVTSAPDNKPAKSTAPPAHVTSPDTPAPGSTAPPAHGVTS
             APNRPALGSTAPPAHVTSPASGASGASTLVHNTSARATTPAPSKETPSIDSH
             SDPTTLIASSTKTDASSTHSTVPLTSSNSTSPQLSTGVSPFLSPHISNLQFNS
             SLSDPSTDYVDELQDISMFLQIVKGQFLGISNKKRPGSVVQLTLAPREGTIN
             LAVDETQNYKTEAASRYNLITSDVSVDPVPPFSAQSGAGVGMGIALIVLCVLA
             LAIVLIALAVCCCRKNYGOLDIFPARDTYHMSERYPIYTHHGRVYVPSSTDRSPYE
             KVASGNGSSLSVTNPVAATASNL"
             73..135
             /gene="MCNAA"
             136..1497
             /gene="MCNAA"
             /product="polymorphic epithelial mucin"
             1783..1788
             /gene="MCNAA"

ORIGIN
Query Match      71.7%; Score 43; DB 9; Length 1804;
Best Local Similarity 83.1%; Pred. No. 0.0078;
Matches          49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Cy              1 GGTTCGACGCGCCCGCTGCTACGCTGTAACATCCGCCCGGATACGACGCGCGCC 59
Db              457 GGCTTCACCGCGCCCGCCGACGCTGTCACCTCGGCCCGGACACGACGCGCGCC 515

RESULT 12
LOCUS       AF423030             1834 bp    mRNA             linear      SYN 10-JUL-2003
DEFINITION   Synthetic construct Homo sapiens mucin variant MUC-1CT80 (MUC1)
ACCESSION    AF423030
VERSION      AF423030.1   GI:19338619
KEYWORDS     synthetic construct
SOURCE       synthetic construct
ORGANISM     artificial sequences.
REFERENCE    1
AUTHORS      Hinojosa-Kurtzberg,A.M., Johansson,M.E., Madsen,C.S., Hansson,G.C.
             and Gendler,S.J.

```

TITLE Novel MUC1 splice variants contribute to mucin overexpression in C57BL/6 mice
JOURNAL Am. J. Physiol. Gastrointest. Liver Physiol. 284 (5), G853-G862 (2003)
MEDLINE 22570517
PUBMED 12529261
REFERENCE 2 (bases 1 to 1834)
AUTHORS Hinojosa-Kurtzberg, A.M. and Gendler, S.J.
TITLE Direct Substitution
JOURNAL Submitted (21-SEP-2001) Biochemistry, Mayo Clinic, 13400 E. Shea Boulevard, Scottsdale, AZ 85259, USA
FEATURES
source
1. .1834
/organism="synthetic construct"
/mol_type="mRNA"
/db_xref="taxon:32630"
/note="Homo sapiens gene in transgenic Mus musculus C57BL/6; isolated from intestinal mucosa"
1. .1834
/gene="MUC1"
1. .1452
/gene="MUC1"
/note="alternatively spliced; contains exon 6a resulting in variant hydrophilic tail of carboxy-terminal domain; lacks sites for beta-catenin and Gb2 interactions; derived from Homo sapiens"
/codon_start=1
/transl_table=11
/product="mucin variant MUC-1CT80"
/protein_id="AAL86734.1"
/db_xref="GI:19338620"
/translation="MTPTGSPFPLLLTLTVTVGSGASTPGGEKSTARRSS VPSSTKNAVSKTSSVLSHSGSGSSTTCGGDVTAPRTBPASGAATGCDVTSVP VADNPALGSTPAPHVNTAPDNKPAAGSTAPAPADTRPAPAGSTAPAPAGVTS APDRPALGSTAPVNTAPVNTASGASGASTLVHNGTASARATTPASKSTPISPSH SDPTTLASHTKTDASTHSTVPLPLTSMNSTPQSLTGVSPFLSPHINLQFNS SLEPSTDYOELODISEMPLQIKGGFLLSNKPRGSGVVOLTLAFREGTINV HDVTFQNOYKTEAASRNLTSDVSVSVPFPFASQASGAPGNGIALVVCVLVA LAIVTLAVALQCRKNTGQLDIFPARITPMSBYPTVTHGRIVPSPSTRSPYE KRLPTGQKQKRWLGMDSGGTWTQKRAMKR"
ORIGIN
Query Match 71.7%; Score 43; DB 12; Length 1834;
Best Local Similarity 83.1%; Pred. No. 0.0078;
Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
Qy 1 GGTTCAGCGCGCCCGCTGCTCAGGTGTACATCCCGCGATACAGACCGGCCCC 59
Db 385 GGCTCCACCGCGCCCGCTGCTCAGGTGTACATCCCGCGATACAGACCGGCCCC 443
RESULT 13
AX959918 1835 bp DNA linear PAT 14-JAN-2004
LOCUS AX959918
DEFINITION Sequence 23 from Patent WO03100060.
ACCESSION AX959918
VERSION AX959918.1 GI:40880145
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Burden, N.G., Ellis, J.H. and Hamblin, P.A.
TITLE Muc-1 antigen with reduced number of vnter repeat units
JOURNAL Patent: WO 03100060-A 23 04-DEC-2003;
GLAXO GROUP LIMITED (GB)
FEATURES
source
1. .1835
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN

Query Match 71.7%; Score 43; DB 6; Length 1835;
Best Local Similarity 83.1%; Pred. No. 0.0078;
Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
Qy 1 GGTTCAGCGCGCCCGCTGCTCAGGTGTACATCCCGCGATACAGACCGGCCCC 59
Db 462 GGCTCCACCGCGCCCGCTGCTCAGGTGTACATCCCGCGATACAGACCGGCCCC 520
RESULT 14
AX959916 2135 bp DNA linear PAT 14-JAN-2004
LOCUS AX959916
DEFINITION Sequence 21 from Patent WO03100060.
ACCESSION AX959916
VERSION AX959916.1 GI:40880144
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Burden, N.G., Ellis, J.H. and Hamblin, P.A.
TITLE Muc-1 antigen with reduced number of vnter repeat units
JOURNAL Patent: WO 03100060-A 21 04-DEC-2003;
GLAXO GROUP LIMITED (GB)
FEATURES
source
1. .2135
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 71.7%; Score 43; DB 6; Length 2135;
Best Local Similarity 83.1%; Pred. No. 0.0075;
Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
Qy 1 GGTTCAGCGCGCCCGCTGCTCAGGTGTACATCCCGCGATACAGACCGGCCCC 59
Db 462 GGCTCCACCGCGCCCGCTGCTCAGGTGTACATCCCGCGATACAGACCGGCCCC 520
RESULT 15
CQ834017 4139 bp DNA linear PAT 29-JUL-2004
LOCUS CQ834017
DEFINITION Sequence 53 from Patent EPI439393.
ACCESSION CQ834017
VERSION CQ834017.1 GI:50833622
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Asle, J.H., Boardman, L.A., Bugart, L.J., Burgess, C.C., Carino, T.J., Diwedi, P., Huntress, M., Johnson, K.A., Lewis, M.E., Matmonis, P.J., Myerow, S.H., Brown-Shimer, S.L., Thibagalingam, A., Thibodeau, S.N. and Molino, G.A.
TITLE Detection methods using TIMP 1 for colon cancer diagnosis
JOURNAL Patent: EP 1439393-A 53 21-JUL-2004;
Bayer Healthcare LLC (US); MAYO FOUNDATION FOR MEDICAL EDUCATION AND RESEARCH (US)
FEATURES
source
1. .4139
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 71.7%; Score 43; DB 6; Length 4139;
Best Local Similarity 83.1%; Pred. No. 0.0065;
Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Oy 1 GGTTGACGAGCCCCCTGCTCACGGTTAACATCCGCCCCGAGATACAGACCGGCCCC 59
Db 458 GGCTCCACCGCCCCCCAGCCCAAGGTGTCACTCCGCCCCGAGACACAGCCCGGCCCC 516

Search completed: January 15, 2005, 22:36:52
Job time : 722.8 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2005, 22:37:00 : Search time 183.4 Seconds
(without alignments)
1879.789 Million cell updates/sec

Title: US-10-057-136-6

Perfect score: 60

Sequence: 1 gggatcaacagctccctccgcg.....cagatattcccccagctcca 60

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4300275 seqs, 2872944193 residues

Total number of hits satisfying chosen parameters: 8600550

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US07_PUB_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_PUB_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_PUB_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_PUB_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_PUB_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10_PUB_PUB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US11_PUB_PUB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US11_PUB_PUB.seq:*
21: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	60	100.0	60	US-10-057-136-6	Sequence 6, Appli
2	60	100.0	2297	US-10-406-317-41	Sequence 41, Appli
3	36.6	61.0	60	US-10-057-136-11	Sequence 11, Appli
4	36	60.0	60	US-10-057-136-10	Sequence 10, Appli
5	31.8	53.0	60	US-10-057-136-5	Sequence 5, Appli
6	30.2	50.3	60	US-10-057-136-7	Sequence 7, Appli
7	30	50.0	60	US-10-057-136-8	Sequence 8, Appli
8	30	50.0	60	US-10-057-136-9	Sequence 9, Appli
9	29.6	49.3	93	US-10-471-607-4	Sequence 4, Appli
10	29.6	49.3	157	US-10-471-607-6	Sequence 6, Appli
11	29	48.3	60	US-10-057-136-12	Sequence 12, Appli
12	28.8	48.0	490	US-10-424-599-89670	Sequence 89670, A

13	28.6	47.7	519	US-10-335-394-42	Sequence 42, Appli
14	28.6	47.7	534	US-10-335-394-46	Sequence 46, Appli
15	27.8	46.3	697	US-09-770-149-278	Sequence 278, Appli
16	27.6	46.0	1065	US-10-425-115-90627	Sequence 90627, A
17	27.6	46.0	2469	US-10-392-810-1	Sequence 1, Appli
18	27.4	45.7	60	US-10-057-136-4	Sequence 4, Appli
19	27.4	45.7	435	US-09-918-995-28868	Sequence 28868, A
20	27.2	45.3	1350	US-10-437-963-42569	Sequence 42569, A
21	27	45.0	60	US-10-057-136-2	Sequence 2, Appli
22	27	45.0	1424	US-10-447-839A-75	Sequence 75, Appli
23	27	45.0	1428	US-10-447-839A-20	Sequence 20, Appli
24	27	45.0	1527	US-10-057-136-19	Sequence 19, Appli
25	27	45.0	1799	US-10-447-839A-19	Sequence 19, Appli
26	27	45.0	1804	US-09-964-824A-573	Sequence 573, Appli
27	27	45.0	1804	US-10-029-517-17	Sequence 17, Appli
28	27	45.0	1804	US-10-717-557-30	Sequence 30, Appli
29	27	45.0	1804	US-10-775-920-84	Sequence 84, Appli
30	27	45.0	1823	US-10-101-510-339	Sequence 339, Appli
31	27	45.0	4139	US-09-964-824A-105	Sequence 105, Appli
32	27	45.0	4139	US-09-964-824A-578	Sequence 578, Appli
33	27	45.0	4139	US-09-864-864-334	Sequence 334, Appli
34	27	45.0	4139	US-09-880-107-2121	Sequence 2121, Appli
35	27	45.0	4139	US-09-968-007A-751	Sequence 751, Appli
36	27	45.0	4139	US-10-171-311-157	Sequence 157, Appli
37	27	45.0	4139	US-10-177-293-310	Sequence 310, Appli
38	27	45.0	4139	US-10-440-464-155	Sequence 155, Appli
39	27	45.0	4139	US-10-734-564-53	Sequence 53, Appli
40	27	45.0	4139	US-10-775-920-80	Sequence 80, Appli
41	27	45.0	4139	US-10-775-920-85	Sequence 85, Appli
42	27	45.0	8181	US-10-447-839A-18	Sequence 18, Appli
43	27	45.0	8186	US-10-247-703-23	Sequence 23, Appli
44	27	45.0	8186	US-10-029-517-19	Sequence 19, Appli
45	26.8	44.7	1826	US-09-938-842A-1210	Sequence 1210, Appli

ALIGNMENTS

RESULT 1
US-10-057-136-6
Sequence 6, Application US/10057136
Publication No. US20030021770A1
GENERAL INFORMATION:
APPLICANT: SCHLOM, JEFFREY
APPLICANT: KANTOR, JUDITH
APPLICANT: KUFE, DONALD
APPLICANT: PANICALI, DENNIS
APPLICANT: GRITZ, LINDA
TITLE OF INVENTION: TUMOR-ASSOCIATED ANTIGEN
FILE REFERENCE: 700953/47113C
CURRENT APPLICATION NUMBER: US/10/057,136
CURRENT FILING DATE: 2002-01-25
PRIOR APPLICATION NUMBER: 09/366, 670
PRIOR FILING DATE: 1998-08-03
PRIOR APPLICATION NUMBER: PCT/US98/03693
PRIOR FILING DATE: 1998-02-24
PRIOR APPLICATION NUMBER: 60/038, 253
PRIOR FILING DATE: 1997-02-24
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 6
LENGTH: 60
TYPE: DNA
ORGANISM: Homo sapiens
US-10-057-136-6

Query Match 100.0%; Score 60; DB 14; Length 60;
Best Local Similarity 100.0%; Pred. No. 4.5e-12;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 gggatcaacagctccctccgcttacttgcctcagatcagccagctcca 60
|||||

Db 1 GGGTCAACAGCTCTCCCGCTCATGGGGTACTTCTGCTCCAGATACTCGCCCAAGCTCCA 60

RESULT 2

US-10-406-317-41

/ Sequence 41, Application US/10406317

/ Publication No. US20040019195A1

/ GENERAL INFORMATION:

/ APPLICANT: Schlom, Jeffrey;

/ APPLICANT: Hodge, James;

/ APPLICANT: Panicali, Dennis

/ TITLE OF INVENTION: A recombinant vector expressing multiple constitutulatory

/ FILE REFERENCE: 38163-0189

/ CURRENT APPLICATION NUMBER: US/10/406,317

/ PRIOR FILING DATE: 2003-04-04

/ PRIOR APPLICATION NUMBER: US/09/856,988

/ PRIOR FILING DATE: 2001-05-30

/ PRIOR APPLICATION NUMBER: PCT/US99/26866

/ PRIOR FILING DATE: 1999-11-12

/ PRIOR APPLICATION NUMBER: 60/111,582

/ PRIOR FILING DATE: 1998-12-09

/ NUMBER OF SEQ ID NOS: 41

/ SOFTWARE: PatentIn Ver. 2.1

/ SEQ ID NO 41

/ LENGTH: 2297

/ TYPE: DNA

/ ORGANISM: Artificial Sequence

/ FEATURE:

/ OTHER INFORMATION: Description of Artificial Sequence: VECTOR

US-10-406-317-41

Query Match 100.0%; Score 60; DB 16; Length 2297;

Best Local Similarity 100.0%; Pred.No.7.7e-12;

Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGTCAACAGCTCTCCCGCTCATGGGGTACTTCTGCTCCAGATACTCGCCCAAGCTCCA 60

Db 526 GGGTCAACAGCTCTCCCGCTCATGGGGTACTTCTGCTCCAGATACTCGCCCAAGCTCCA 585

RESULT 3

US-10-057-136-11

/ Sequence 11, Application US/10057136

/ Publication No. US20030021770A1

/ GENERAL INFORMATION:

/ APPLICANT: SCHLOM, JEFFREY

/ APPLICANT: KANTOR, JUDITH

/ APPLICANT: KUFE, DONALD

/ APPLICANT: PANICALI, DENNIS

/ APPLICANT: GRITZ, LINDA

/ TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1

/ FILE REFERENCE: 700953/47113C

/ CURRENT APPLICATION NUMBER: US/10/057,136

/ PRIOR FILING DATE: 2002-01-25

/ PRIOR APPLICATION NUMBER: 09/366,670

/ PRIOR FILING DATE: 1999-08-03

/ PRIOR APPLICATION NUMBER: PCT/US98/03693

/ PRIOR FILING DATE: 1998-02-24

/ PRIOR APPLICATION NUMBER: 60/038,253

/ PRIOR FILING DATE: 1997-02-24

/ NUMBER OF SEQ ID NOS: 20

/ SOFTWARE: PatentIn Ver. 2.1

/ SEQ ID NO 11

/ LENGTH: 60

/ TYPE: DNA

/ ORGANISM: Homo sapiens

US-10-057-136-11

Query Match 61.0%; Score 36.6; DB 14; Length 60;

Best Local Similarity 76.3%; Pred. No. 0.0014;

Matches 45; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

OY 1 GGGTCAACAGCTCTCCCGCTCATGGGGTACTTCTGCTCCAGATACTCGCCCAAGCTCC 59

Db 1 GGTTCACGGCAGCTCTCCAGCAGCAGGAGTCAGCTGTGACCCGACCCCGTCCAGCTCC 59

RESULT 4

US-10-057-136-10

/ Sequence 10, Application US/10057136

/ Publication No. US20030021770A1

/ GENERAL INFORMATION:

/ APPLICANT: SCHLOM, JEFFREY

/ APPLICANT: KANTOR, JUDITH

/ APPLICANT: KUFE, DONALD

/ APPLICANT: PANICALI, DENNIS

/ APPLICANT: GRITZ, LINDA

/ TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1

/ FILE REFERENCE: 700953/47113C

/ CURRENT APPLICATION NUMBER: US/10/057,136

/ PRIOR FILING DATE: 2002-01-25

/ PRIOR APPLICATION NUMBER: 09/366,670

/ PRIOR FILING DATE: 1999-08-03

/ PRIOR APPLICATION NUMBER: PCT/US98/03693

/ PRIOR FILING DATE: 1998-02-24

/ PRIOR APPLICATION NUMBER: 60/038,253

/ PRIOR FILING DATE: 1997-02-24

/ NUMBER OF SEQ ID NOS: 20

/ SOFTWARE: PatentIn Ver. 2.1

/ SEQ ID NO 10

/ LENGTH: 60

/ TYPE: DNA

/ ORGANISM: Homo sapiens

US-10-057-136-10

Query Match 60.0%; Score 36; DB 14; Length 60;

Best Local Similarity 75.0%; Pred.No.0.0023;

Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

OY 1 GGGTCAACAGCTCTCCCGCTCATGGGGTACTTCTGCTCCAGATACTCGCCCAAGCTCCA 60

Db 1 GGGTCAAGCTGCGCCCTCCGGCGCATGGTGTGACCTCAGCTCTCAGACACAGCCAGCCCA 60

RESULT 5

US-10-057-136-5

/ Sequence 5, Application US/10057136

/ Publication No. US20030021770A1

/ GENERAL INFORMATION:

/ APPLICANT: SCHLOM, JEFFREY

/ APPLICANT: KANTOR, JUDITH

/ APPLICANT: KUFE, DONALD

/ APPLICANT: PANICALI, DENNIS

/ APPLICANT: GRITZ, LINDA

/ TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1

/ FILE REFERENCE: 700953/47113C

/ CURRENT APPLICATION NUMBER: US/10/057,136

/ PRIOR FILING DATE: 2002-01-25

/ PRIOR APPLICATION NUMBER: 09/366,670

/ PRIOR FILING DATE: 1999-08-03

/ PRIOR APPLICATION NUMBER: PCT/US98/03693

/ PRIOR FILING DATE: 1998-02-24

/ PRIOR APPLICATION NUMBER: 60/038,253

/ PRIOR FILING DATE: 1997-02-24

/ NUMBER OF SEQ ID NOS: 20

/ SOFTWARE: PatentIn Ver. 2.1

/ SEQ ID NO 5

/ LENGTH: 60

/ TYPE: DNA

/ ORGANISM: Homo sapiens

US-10-057-136-5

LENGTH: 60

OTHER INFORMATION: Artificial epitope construct
US-10-471-607-4

```
Query Match          49.3%; Score 29.6; DB 17; Length 93;
Best Local Similarity 79.5%; Pred. No. 0.5;
Matches 35; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 GGGTCAACAGCTCTCCCGCTCATGGGGTTACTTGTCTCCAGA 44
Db 57 GGGTCAACAGCCCCCGAGCTCATGGTGTCACTCAGCTCCGA 14

RESULT 10
US-10-471-607-6
; Sequence 6, Application US/10471607
; Publication No. US2004015740A1
; GENERAL INFORMATION:
; APPLICANT: The Victoria University of Manchester
; APPLICANT: Benson, Roderick
; TITLE OF INVENTION: Intracellular analysis.
; FILE REFERENCE: P088857PWO
; CURRENT APPLICATION NUMBER: US/10/471,607
; CURRENT FILING DATE: 2003-09-24
; PRIOR APPLICATION NUMBER: GB 0108165.2
; PRIOR FILING DATE: 3001-03-21
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 157
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Artificial epitope construct
US-10-471-607-6

Query Match          49.3%; Score 29.6; DB 17; Length 157;
Best Local Similarity 79.5%; Pred. No. 0.54;
Matches 35; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 GGGTCAACAGCTCTCCCGCTCATGGGGTTACTTGTCTCCAGA 44
Db 101 GGGTCAACAGCCCCCGAGCTCATGGTGTCACTCAGCTCCGA 144

RESULT 11
US-10-057-136-12
; Sequence 12, Application US/10057136
; Publication No. US20030021770A1
; GENERAL INFORMATION:
; APPLICANT: SCHLON, JEFFREY
; APPLICANT: KANTOR, JUDITH
; APPLICANT: KUFE, DONALD
; APPLICANT: PANICALI, DENNIS
; APPLICANT: GRITZ, LINDA
; TITLE OF INVENTION: TUMOR-ASSOCIATED ANTIGEN
; TITLE OF INVENTION: TUMOR-ASSOCIATED ANTIGEN
; FILE REFERENCE: 700953/471136
; CURRENT APPLICATION NUMBER: US/10/057,136
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/366,670
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: PCT/US98/03693
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/038,253
; PRIOR FILING DATE: 1997-02-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-136-12

Query Match          48.3%; Score 29; DB 14; Length 60;
Best Local Similarity 71.7%; Pred. No. 0.77;
```

```
Matches 38; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 7 ACAGTCTCTCCCGCTCATGGGGTTACTTGTCTCAGATATCTGCCAGCTCC 59
Db 7 ACAGGCCACCCGCGACATAGCGGTCAAGCGCTCCGATACGAGACCGGCGCC 59

RESULT 12
US-10-424-599-89670/C
; Sequence 89670, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yinhua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 89670
; LENGTH: 490
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_51983C.1
US-10-424-599-89670

Query Match          48.0%; Score 28.8; DB 16; Length 490;
Best Local Similarity 69.6%; Pred. No. 1.2;
Matches 39; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 5 CAACAGCTCTCTCCCGCTCATGGGGTTACTTGTCTCAGATATCTGCCAGCTCC 60
Db 273 CAGCGGCTTCTTGGGCTTCAGCGGCTTCTTGGGCTTCAGCGGCTTGTGCTGCTCA 218

RESULT 13
US-10-335-394-42
; Sequence 42, Application US/10335394
; Publication No. US20030138452A1
; GENERAL INFORMATION:
; APPLICANT: Kwak, Larry
; APPLICANT: Biraayn, Arya
; TITLE OF INVENTION: METHODS AND COMPOSITIONS OF
; TITLE OF INVENTION: CHEMOKINE-TUMOR ANTIGEN FUSION PROTEINS AS CANCER VACCINES
; FILE REFERENCE: 14014.0316/P
; CURRENT APPLICATION NUMBER: US/10/335,394
; CURRENT FILING DATE: 2002-12-31
; PRIOR APPLICATION NUMBER: US/09/646,028
; PRIOR FILING DATE: 2000-09-12
; PRIOR APPLICATION NUMBER: 60/077,745
; PRIOR FILING DATE: 1998-03-12
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 42
; LENGTH: 519
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of artificial sequence: note=synthetic construct
US-10-335-394-42

Query Match          47.7%; Score 28.6; DB 15; Length 519;
Best Local Similarity 67.8%; Pred. No. 1.5;
Matches 40; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 1 GGGTCAACAGCTCTCCCGCTCATGGGGTTACTTGTCTCAGATATCTGCCAGCTCC 59
Db 355 GGTGTACTTCTGTGGCGCTTCGACGGGTGTACTTGTGCCCCGAGACTGCCAGACCC 413
```



```
RESULT 14
US-10-335-394-46
; Sequence 46, Application US/10335394
; Publication No. US20030138452A1
; GENERAL INFORMATION:
; APPLICANT: Kwak, Larry
; APPLICANT: Biregyn, Arya
; TITLE OF INVENTION: METHODS AND COMPOSITIONS OF
; FILE REFERENCE: 14014.0316/P
; CURRENT APPLICATION NUMBER: US/10/335,394
; CURRENT FILING DATE: 2002-12-31
; PRIOR APPLICATION NUMBER: US/09/646,028
; PRIOR FILING DATE: 2000-09-12
; PRIOR APPLICATION NUMBER: 60/077,745
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 46
; LENGTH: 534
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of artificial sequence:/note=synthetic construct
US-10-335-394-46

Query Match          47.7%; Score 28.6; DB 15; Length 534;
Best Local Similarity 67.8%; Pred. No. 1.5;
Matches 40; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY      1 GGGTCAACAGCTCTCCCGCTCATGGGGTACTTGCTCCAGATACCTGCGCCAGCTCC 59
DB      370 GGTGACTTCTGCGCCCGCTCGACGGTGTACTTCTGCCCCGACACTGCGCCAGCACC 428

RESULT 15
US-09-770-149-278/C
; Sequence 278, Application US/09770149
; Patent No. US2002005963A1
; GENERAL INFORMATION:
; APPLICANT: Gorlach, Jorn
; APPLICANT: An, Yong-Qiang
; APPLICANT: Hamilton, Carol M.
; APPLICANT: Price, Jennifer L.
; APPLICANT: Raines, Tracy M.
; APPLICANT: Yu, Yang
; APPLICANT: Rameaka, Joshua G.
; APPLICANT: Page, Amy
; APPLICANT: Mathew, Abraham V.
; APPLICANT: Ledford, Brooke L.
; APPLICANT: Woessner, Jeffrey P.
; APPLICANT: Haas, William David
; APPLICANT: Garcia, Carlos A.
; APPLICANT: Kricker, Maja
; APPLICANT: Slader, Ted
; APPLICANT: Davis, Keith R.
; APPLICANT: Allen, Keith
; APPLICANT: Hoffman, Neil
; APPLICANT: Hurban, Patrick
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis
; FILE REFERENCE: 2024 (PARA-013PRV)
; CURRENT APPLICATION NUMBER: US/09/770,149
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/178,506
; PRIOR FILING DATE: 2000-01-27
; NUMBER OF SEQ ID NOS: 999
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 278
; LENGTH: 697
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
```

```
US-09-770-149-278

Query Match          46.3%; Score 27.8; DB 9; Length 697;
Best Local Similarity 69.1%; Pred. No. 3;
Matches 38; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY      5 CAACAGCTCTCCCGCTCATGGGGTACTTGCTCCAGATACCTGCGCCAGCTCC 59
DB      373 CAAGTCTTCTGCGCCGCTGAAGCGATGCTGCAGCTGCACTGCAAGCTGC 319

Search completed: January 16, 2005, 09:30:24
Job time : 183.4 secs
```

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2005, 19:50:13 : Search time 35.2 Seconds
(without alignments)
1211.572 Million cell updates/sec

Title: US-10-057-136-6
Perfect score: 60
Sequence: 1 gggcacaagcgtccctccgc.....cagatccgcgcagctcca 60

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:
1: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/prodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
4: /cgn2_6/prodata/1/ina/5B_COMB.seq:*
5: /cgn2_6/prodata/1/ina/5C_COMB.seq:*
6: /cgn2_6/prodata/1/ina/Backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	28.6	47.7	519	US-09-646-028-42	Sequence 42, Appl
2	28.6	47.7	534	US-09-646-028-46	Sequence 46, Appl
3	27.6	46.0	303	US-09-489-039A-2070	Sequence 2070, Ap
4	27	45.0	1804	US-10-029-517-17	Sequence 17, Appl
5	27	45.0	8186	US-10-029-517-19	Sequence 19, Appl
6	26.4	44.0	2301	US-08-306-691B-23	Sequence 23, Appl
7	26.4	44.0	2301	US-09-167-206-3	Sequence 3, Appl
8	26.4	44.0	2301	PCT-US93-06251-78	Sequence 78, Appl
9	26.4	44.0	16442	US-08-781-891-208	Sequence 208, App
10	26.4	44.0	16442	US-09-618-166-208	Sequence 208, App
11	26.2	43.7	1013	US-09-270-767-4588	Sequence 4588, Ap
12	26.2	43.7	1013	US-09-270-767-19870	Sequence 19870, A
13	25.6	42.7	205	US-09-513-999C-13924	Sequence 13924, A
14	25.6	42.7	453	US-09-221-017B-434	Sequence 434, App
15	25.6	42.7	518	US-10-029-517-101	Sequence 101, App
16	25.6	42.7	981	US-10-029-517-16	Sequence 16, Appl
17	25.6	42.7	1721	US-10-029-517-3	Sequence 3, Appl
18	25.6	42.7	1952	US-09-520-781-23	Sequence 23, Appl
19	25.6	42.7	2010	US-09-520-781-21	Sequence 21, Appl
20	25.6	42.7	3343	US-10-029-517-102	Sequence 102, App
21	25.4	42.3	572	US-10-029-517-18	Sequence 18, Appl
22	25.4	42.3	1770	US-09-489-039A-3639	Sequence 3639, Ap
23	25.4	42.3	1230025	US-09-198-452A-1	Sequence 1, Appl
24	25.2	42.0	418	US-09-270-767-25147	Sequence 25147, A
25	25.2	42.0	486	US-09-270-767-9852	Sequence 9852, Ap
26	25	41.7	2304	US-09-799-451-223	Sequence 223, App
27	25	41.7	6972	US-09-595-684B-38	Sequence 38, Appl

C	28	25	41.7	8309	4	US-09-620-312D-1083	Sequence 1083, Ap
	29	24.8	41.3	994	6	5202236-39	Patent No. 5202236
	30	24.8	41.3	1236	6	5202236-4	Patent No. 5202236
	31	24.8	41.3	1607	3	US-08-753-007A-3	Sequence 3, Appl
	32	24.8	41.3	1607	3	US-09-398-496-3	Sequence 3, Appl
	33	24.8	41.3	1926	4	US-09-249-585A-2	Sequence 2, Appl
	34	24.8	41.3	1926	4	US-09-410-399-3	Sequence 3, Appl
	35	24.8	41.3	2467	3	US-08-753-007A-1	Sequence 1, Appl
	36	24.8	41.3	2467	3	US-09-398-496-1	Sequence 1, Appl
	37	24.8	41.3	2580	3	US-09-050-863-2	Sequence 2, Appl
	38	24.8	41.3	2580	3	US-09-359-081-2	Sequence 2, Appl
	39	24.8	41.3	5452	2	US-09-130-114-1	Sequence 1, Appl
	40	24.8	41.3	8705	4	US-09-647-344A-14	Sequence 14, Appl
	41	24.8	41.3	9600	4	US-08-910-647-1	Sequence 1, Appl
	42	24.8	41.3	9600	4	US-09-620-925-1	Sequence 1, Appl
	43	24.8	41.3	10596	1	US-07-884-811-15	Sequence 15, Appl
	44	24.8	41.3	10596	1	US-07-885-971-15	Sequence 15, Appl
	45	24.8	41.3	10596	1	US-08-087-783A-15	Sequence 15, Appl

ALIGNMENTS

```

RESULT 1
US-09-646-028-42
; Sequence 42, Application US/09646028
; Patent No. 6562347
; GENERAL INFORMATION:
; APPLICANT: Kwak, Larry
; APPLICANT: Birsayn, Arya
; TITLE OF INVENTION: METHODS AND COMPOSITIONS OF
; FILE REFERENCE: 14014.0316/P
; CURRENT APPLICATION NUMBER: US/09/646,028
; PRIOR FILING DATE: 2000-09-12
; PRIOR APPLICATION NUMBER: 60/077,745
; PRIOR FILING DATE: 1998-03-12
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 42
; LENGTH: 519
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of artificial sequence://note=synthetic construct
US-09-646-028-42

Query Match          47.7% Score 28.6; DB 4; Length 519;
Best Local Similarity 67.8% Pred. No. 0.67;
Matches 40; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY      1 GGGTCAACAGCTCCCGCGCTCATGGGTTACTTGTGTCAGATACTGCCAGCTCC 59
DB      355 GGTGTAACCTTTCGCCCGCTCATGCGGTGTAACCTTGTGCGCCGACACTGCCAGACCC 413

RESULT 2
US-09-646-028-46
; Sequence 46, Application US/09646028
; Patent No. 6562347
; GENERAL INFORMATION:
; APPLICANT: Kwak, Larry
; APPLICANT: Birsayn, Arya
; TITLE OF INVENTION: METHODS AND COMPOSITIONS OF
; FILE REFERENCE: 14014.0316/P
; CURRENT APPLICATION NUMBER: US/09/646,028
; PRIOR FILING DATE: 2000-09-12
; PRIOR APPLICATION NUMBER: 60/077,745
; PRIOR FILING DATE: 1998-03-12
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 46

```

```

; LENGTH: 534
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of artificial sequence:/note=synthetic construct
US-09-646-028-46
```

```

Query Match
Best Local Similarity 47.7%; Score 28.6; DB 4; Length 534;
Matches 40; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
```

```

QY 1 GGGTCAACAGCTCTCCCGCTCATGGGATTACTCTGCTCCAGATCTGCCAGCTCC 59
DB 370 GGTTAACTTCTGCCCCCTCGACGGTGAATCTTGCCCCGACACTGCGCAGCACC 428
```

RESULT 3

```

US-09-489-039A-2070/c
; Sequence 2070, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; NUMBER OF SEQ ID NOS: 1999-01-29
; SEQ ID NO 2070
; LENGTH: 303
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-2070
```

```

Query Match
Best Local Similarity 46.0%; Score 27.6; DB 4; Length 303;
Matches 39; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
```

```

QY 1 GGGTCAACAGCTCTCCCGCTCATGGGATTACTCTGCTCCAGATCTGCCAGCTC 58
DB 157 GGGTCAGCTCGCATCGAGCGCCCTGGGGAATTTCTTCTCAATCTCTACCCAGATC 100
```

RESULT 4

```

US-10-029-517-17
; Sequence 17, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Doble
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517
; PRIOR FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 17
; LENGTH: 1804
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (73)...(1500)
US-10-029-517-17
```

```

Query Match
Best Local Similarity 45.0%; Score 27; DB 4; Length 1804;
Matches 39; Conservative 0; Mismatches 20; Indels 0; Gaps 0;
```

```

QY 1 GGGTCAACAGCTCTCCCGCTCATGGGATTACTCTGCTCCAGATCTGCCAGCTCC 59
DB 457 GGCTCACCGCGCCCCCGACGCGTGTCACTGCGCCCCGACACAGGCGGCCCC 515
```

RESULT 5

```

US-10-029-517-19
; Sequence 19, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Doble
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517
; PRIOR FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 19
; LENGTH: 8186
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 6899
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 7155
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 7184
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 7957
; OTHER INFORMATION: unknown
; NAME/KEY: intron
; LOCATION: (2997)...(3498)
; OTHER INFORMATION: intron 1
; NAME/KEY: intron:exon junction
; LOCATION: (3498)...(3499)
; OTHER INFORMATION: intron 1:exon 2
; NAME/KEY: exon
; LOCATION: (3508)...(3599)
; OTHER INFORMATION: exon 2d
; NAME/KEY: exon:intron junction
; LOCATION: (3982)...(3983)
; OTHER INFORMATION: exon 2a:intron 2a
; NAME/KEY: intron:exon junction
; LOCATION: (4205)...(4206)
; OTHER INFORMATION: intron 2c:exon 3c
; NAME/KEY: intron:exon junction
; LOCATION: (4259)...(4260)
; OTHER INFORMATION: intron 2d:exon 3d
; NAME/KEY: exon
; LOCATION: (4260)...(4328)
; OTHER INFORMATION: exon 3d
; NAME/KEY: intron:exon junction
; LOCATION: (4632)...(4633)
; OTHER INFORMATION: intron 3:exon 4
; NAME/KEY: exon
; LOCATION: (4914)...(5035)
; OTHER INFORMATION: exon 5
; NAME/KEY: intron
; LOCATION: (5266)...(6293)
; OTHER INFORMATION: intron 6
US-10-029-517-19
```

```

Query Match
Best Local Similarity 45.0%; Score 27; DB 4; Length 8186;
Matches 39; Conservative 0; Mismatches 20; Indels 0; Gaps 0;
```

```

QY 1 GGGTCAACAGCTCTCCCGCTCATGGGATTACTCTGCTCCAGATCTGCCAGCTCC 59
DB 3825 GGCTCACCGCGCCCCCGACGCGTGTCACTGCGCCCCGACACAGGCGGCCCC 3883
```

RESULT 6

```

US-08-306-691B-23/c
; Sequence 23, Application US/08306691B
; Patent No. 5734039
; GENERAL INFORMATION:
; APPLICANT: Calabretta, Bruno
; APPLICANT: Skoraki, Tomasz
; TITLE OF INVENTION: ANTISENSE
; TITLE OF INVENTION: OLIGONUCLEOTIDES TARGETING COOPERATING ONCOGENES
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seidel, Gonda, Lavorgna & Monaco, P.C.
; STREET: Two Penn Center, Suite 1800
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/306,691B
; FILING DATE: September 15, 1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 8321-8
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; TELEX: No. 5734039e
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2301 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
;
US-08-306-691B-23
;
Query Match 44.0%; Score 26.4; DB 1; Length 2301;
Best Local Similarity 69.2%; Pred. No. 5.5'
Matches 36; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
;
OY 8 CAGCTCTCCGCGCTCATGAGGGTACTTGTGCTCCAGATATCGCCAGCTCC 59
DB 85 CTGCTCTCTCCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCC 34
;
RESULT 7
; Sequence 3, Application US/09167206A
; Patent No. 6476193
; GENERAL INFORMATION:
; APPLICANT: Nandabalan, Krishnan
; APPLICANT: Schulz, Vincent P.
; APPLICANT: Yang, WeiJa
; TITLE OF INVENTION: NIK1 PROTEIN AND NIK1 PROTEIN COMPLEXES
; FILE REFERENCE: 15966-521 NIK1 protein complexes
; CURRENT APPLICATION NUMBER: US/09/167,206A
; CURRENT FILING DATE: 1998-10-06
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0.
; SEQ ID NO 3
; LENGTH: 2301
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS

```

[illegible]

Db 13 CAGCTCAGCTCAGCTCCTGGGGTTCTTCTTACCAAGTCTGTTCAGTACC 67

RESULT 13

US-09-513-999C-13924/C
Sequence 13924, Application US/09513999C
Patent No. 6783961
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Duclert, A.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
Patent No. 6783961
FILE REFERENCE: 59, US2, REG
CURRENT APPLICATION NUMBER: US/09/513, 999C
CURRENT FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/122,487
PRIOR FILING DATE: 1999-02-26
NUMBER OF SEQ ID NOS: 36681
SOFTWARE: Patent.pm
SEQ ID NO 13924
LENGTH: 205
TYPE: DNA
ORGANISM: Homo sapiens
US-09-513-999C-13924

Query Match 42.7%; Score 25.6; DB 4; Length 205;
Best Local Similarity 66.1%; Pred. No. 5.8;

Matches 37; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 5 CAACGCTCTCCGCTCATGGGTTACTTCTGCTCCAGATCTCGCCAGCTCCA 60
Db 138 CAGGCGTCTCCGCTCCAGGCGTCTGCGCTCCAGGCGTCTCGCGCTCCA 83

RESULT 14

US-09-221-017B-434/C
Sequence 434, Application US/09221017B
Patent No. 6444799
GENERAL INFORMATION:
APPLICANT: Rose, Bruce C.
TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
NUMBER OF SEQUENCES: 1120
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 755 PAGE MILL ROAD
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows
SOFTWARE: PasteSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/221, 017B
FILING DATE: 23-DEC-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PP1182
FILING DATE: 31-DEC-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PP1546
FILING DATE: 30-JAN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PP2911
FILING DATE: 09-APR-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU98/01023
FILING DATE: 10-DEC-1998
ATTORNEY/AGENT INFORMATION:

NAME: Monroy, Gladys H
REGISTRATION NUMBER: 32,430
REFERENCE/DOCKET NUMBER: 27340-20021.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-813-5600
TELEFAX: 650-494-0792
TELEX: 706141

INFORMATION FOR SEQ ID NO: 434:
SEQUENCE CHARACTERISTICS:

LENGTH: 453 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)

HYPOTHEICAL: NO
ANTI-SENSE: UNKNOWN
ORIGINAL SOURCE:
ORGANISM: PORYPHYROMONAS GINGIVALIS
FEATURE:
NAME/KEY: misc feature
LOCATION: 1...453
US-09-221-017B-434

Query Match 42.7%; Score 25.6; DB 4; Length 453;
Best Local Similarity 70.8%; Pred. No. 7;

Matches 34; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 9 AACTCTCCCGCTCATGGGTTACTTCTGCTCCAGATCTCGCCAGC 56
Db 420 AACTCTCTCCGCTCATGGGTTACTTCTGCTCCAGATCTCGCCAGC 373

RESULT 15

US-10-029-517-101
Sequence 101, Application US/10029517
Patent No. 6716627
GENERAL INFORMATION:
APPLICANT: Kenneth W. Dobie
APPLICANT: Susan J. Myers
TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
FILE REFERENCE: RTS-0352
CURRENT APPLICATION NUMBER: US/10/029,517
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 107
SEQ ID NO 101
LENGTH: 518
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
US-10-029-517-101

Query Match 42.7%; Score 25.6; DB 4; Length 518;
Best Local Similarity 66.1%; Pred. No. 7.2;

Matches 37; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 1 GGCTCAACAGCTCTCCGCTCATGGGTTACTTCTGCTCCAGATCTCGCCAGC 56
Db 462 GGCTCAACAGCTCTCCGCTCATGGGTTACTTCTGCTCCAGATCTCGCCAGC 517

Search completed: January 16, 2005, 03:01:28
Job time : 37.2 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2005, 19:43:55 : Search time 1550.6 Seconds
(without alignments)
1410.024 Million cell updates/sec

Title: US-10-057-136-6

Perfect score: 60

Sequence: 1 GGGTCACAGCTCCCTCCGCGC.....CAGATACCTGCCAGCTCCA 60

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	32.6	54.3	603	7	CF809795 Lr_LCIED
C 2	32	53.3	418	8	AZ943545 2M0204E18
C 3	30.4	50.7	652	9	CNS0303UR Tetradon
C 4	30.4	50.7	983	9	CNS0303UR Tetradon
C 5	29.8	49.7	1101	9	CNS0507K8 Tetradon
C 6	29.6	49.3	413	4	BJ511928 B511928
C 7	29.6	49.3	568	4	BJ498303 B498303
C 8	29.6	49.3	584	4	BJ500902 B500902
C 9	29.6	49.3	612	4	BJ493595 B493595
C 10	29.6	49.3	652	4	BJ014441 B014441
C 11	29.6	49.3	652	4	BJ490836 B490836
C 12	29.6	49.3	669	4	BJ504955 B504955
C 13	29.6	49.3	686	4	BJ028611 B028611
C 14	29.6	49.3	697	4	BJ710454 B710454
C 15	29.6	49.3	730	4	BJ508359 B508359
C 16	29.6	49.3	739	4	BJ516738 B516738
C 17	29.6	49.3	798	4	BJ721686 B721686
C 18	29.6	49.0	644	3	CNS0408AA Arabidops
C 19	29.2	48.7	650	6	CA084017 SCCCL400
C 20	29.2	48.7	684	6	CA083318 SCEPM201
C 21	29.2	48.7	731	6	CA289703 SCAGFL800
C 22	29.2	48.7	871	6	CL566510 OB_Ba003
C 23	29	48.3	209	9	CE134399 tigr-gss-
C 24	29	48.3	341	1	AA591111 vml2c06.r

C 25	29	48.3	345	1	AA967806
C 26	29	48.3	380	7	W65797 mel1402.r1
C 27	29	48.3	393	1	AA153891
C 28	29	48.3	436	2	BB863659
C 29	29	48.3	454	1	AI892918
C 30	29	48.3	523	5	BQ745314 UI-M-ERO-
C 31	29	48.3	838	6	CB201841 AGENCOURT
C 32	29	48.3	874	4	B1525902 602924866
C 33	29	48.3	893	5	BQ939292 AGENCOURT
C 34	29	48.3	1295	9	AY411302 Mus muscu
C 35	28.8	48.0	684	9	CNS041N8
C 36	28.8	48.0	789	9	CNS02HOS
C 37	28.8	48.0	809	9	CNS01V6B
C 38	28.8	48.0	844	9	AG551682 Mus muscu
C 39	28.6	47.7	330	1	AI925867 w020004.x
C 40	28.6	47.7	430	5	BY258590 BY258590
C 41	28.6	47.7	471	7	H19186
C 42	28.6	47.7	1113	5	BU148487
C 43	28.6	47.7	1234	5	BQ936898
C 44	28.6	47.7	1262	5	BQ935496
C 45	28.6	47.7	1262	5	BQ935496

ALIGNMENTS

RESULT 1
CF809795/c
LOCUS
DEFINITION
Lr_LCIED 31G07 SAC Lumbriacus rubellus Late Cocoon Library 1
VERSION
CF809795
KEYWORDS
CF809795.1 GI:37998126
SOURCE
Lumbriacus rubellus (humus earthworm)
ORGANISM
Lumbriacus rubellus
Bukaryota; Metazoa; Annelida; Clitellata; Oligochaeta; Haplotaxida;
Lumbriacina; Lumbriacidae; Lumbriacus.
REFERENCE
1 (bases 1 to 603)
Jones, M., Chaeley, J., Hedley, B.A., Morgan, J.C., Sturzenbaum, S.,
Kille, P. and Blaxter, M.
The Lumbriacus rubellus EST program - Sequences from a cocoon
library
JOURNAL
Unpublished (2003)
COMMENT
Contact: Blaxter, M.
Institute of Cell, Animal and Population Biology
University of Edinburgh
Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9
3JT, UK.
Tel: +44 131 650 6760
Fax: +44 131 670 5450
Email: mark.blaxter@ed.ac.uk
The library was prepared using protocols given by the supplier
(Stratagene).
PCR Primers
FORWARD: M13R(AGCGATACATTTGCACACAGA)
BACKWARD: T7PL(CTGACTATAGGCGAATTGG)
Plate: 31 row: G column: 07
Seq primer: SAC(GGGAACAAGCTGAG)
High quality sequence scop: 504.
location/Qualifiers
1..603
/organism="Lumbriacus rubellus"
/mol_type="mRNA"
/db_xref="taxon:35632"
/clone="Lr_LCIED 31G07"
/tissue_type="Whole worm"
/dev_stage="Late Cocoon"
/clone_lib="Lumbriacus rubellus Late Cocoon Library 1"
/note="Vector: pBluescript II SK+; The library was
prepared using protocols given by the supplier
(Stratagene)."

ORIGIN

```

Only Match: 53.3%; Score 32; DB 8; Length 418;
Best Local Similarity 73.2%; Pred. No. 6.8;
Matches 41; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 5 CACACGCTCTCCGCTCATGGGGTACTCTCTGCTCAGATATCTGCCCGACGTCCA 60
Db 53 CAGCTCTCTCTCTGCTCTGCTCTCTCTCTCCACGCTCAGCTATCTCTCAGTCCA 108

```

RESULT 3	
CNS03USR	
LOCUS	652 bp DNA linear GSS 01-SEP-2000
DEFINITION	Tetradon nigroviridis genome survey sequence.T7 end of clone 059b19 of library G from Tetradon nigroviridis, genomic survey sequence.
ACCESSION	AL261540
VERSION	AL261540.1 GI:7983165
KEYWORDS	GSS; genome survey sequence.
SOURCE	Tetradon nigroviridis
ORGANISM	Tetradon nigroviridis
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthopterygii; Acanthopterygii; Percormorpha; Tetraodontiformes;

TITLE
Bernot, A., Fitzames, C., Wincker, P., Brottier, P., Quettier, F.,
Sautin, W. and Weissenbach, J.
Estimate of human gene number provided by genome-wide analysis
using Tetraodon nigroviridis DNA sequence

PUBMED REFERENCE AUTHORS	TITLE	JOURNAL MEDLINE
10835645	2	
	Roest Crollings,H., Jallion,O., Dasilva,C., Ozouf-Costaz,C., Rizmes,C., Fischer,C., Bonneau,L., Billault,A., Queller,F., Sautin,W., Bernot,A. and Weissenbach,J. Characterization and repeat analysis of the compact genome of the freshwater puffertish Tetraodon lineolus	Genome Res. 10 (7), 939-949 (2000)
	20359837	
	10899143	
	3 (bases 1 to 652)	
	Genoscope.	

COMMENT
This sequence is a single read and was generated as part of a larger genome clone-and-sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at <http://www.genoscope.cns.fr/tetraodon>.

FEATURES

source

1..652
Location/Qualifiers

/organism="Tetraodon nigroviridis"
/mol_type="genomic DNA"
/db_xref="taxon:99883"
/clone="059N19"
/clone_id="G"
/note="Genoscope sequence ID : C0BG059CG10LPI-end : T7"

ORIGIN

Query Match	Best Local Similarity	Matches	Conservative	Mismatches	Indels	Gaps
5 CAACAGCTCCCTCCGGCTCATGSGGTAACTTCCTCCACAGATACTGCCCAAGCTCCA 60	71.4%	40;	0;	16;	0;	0;
Db 136 CACCTGCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCA 191						

RESULT 4

CNS03W76/c 983 bp DNA 1linear GSS 01-SEP-2000
 LOCUS CNS03W76
 DEFINITION Tetraodon nigroviridis genome survey sequence T7 end of clone 063111 of library G from Tetraodon nigroviridis, genomic survey sequence.
 ACCESSION AL263355.1 GI:7985013
 VERSION AL263355
 KEYWORDS GSS: genome survey sequence.
 SOURCE Tetraodon nigroviridis
 ORGANISM Tetraodon nigroviridis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodontidae; Tetraodon.

REFERENCE 1
 Roest Crolius, H., Jallion, O., Dasilva, C., Bouneau, L., Fisher, C., Bernot, A., Fizames, C., Wincker, P., Brottier, P., Quetier, F., Saurin, W., and Weissenbach, J.
 Estimate of human gene number provided by genome-wide analysis using Tetraodon nigroviridis DNA sequence
 Nat. Genet. 25 (2), 235-238 (2000)

JOURNAL MEDLINE 20296633
 PUBMED 10835645
 REFERENCE 2
 Roest Crolius, H., Jallion, O., Dasilva, C., Ozouf-Costaz, C., Fizames, C., Fischer, C., Bouneau, L., Billault, A., Quetier, F., Saurin, W., Bernot, A., and Weissenbach, J.
 Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis
 Genome Res. 10 (7), 939-949 (2000)

JOURNAL MEDLINE 20359837
 PUBMED 10899143
 REFERENCE 3 (bases 1 to 983)
 Genoscope.
 Direct Submission
 Submitted (12-APR-2000) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
 This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at <http://www.genoscope.cns.fr/Tetraodon>.

FEATURES
 .source
 1.983
 /organism="Tetraodon nigroviridis"
 /mol_type="genomic DNA"
 /db_xref="taxon:99883"
 /clone="063111"
 /clone_1ib="G"
 /note="Genoscope sequence ID : COB063CF06LPI-end : T7"

ORIGIN
 Query Match 50.7%; Score 30.4; DB 9; Length 983;
 Best Local Similarity 71.4%; Pred. No. 25;
 Matches 40; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 5 CAACAGCTCTCTCCGCTCATGGGTTACTTGTGCTCCAGATCTGCCAGCTTCA 60
 Db 862 CACCTGCTCA 807

RESULT 5
 LOCUS CNS057K8/c 1101 bp DNA 1linear GSS 26-JUL-2000
 DEFINITION Tetraodon nigroviridis genome survey sequence SP6 end of clone 002K02 of library B from Tetraodon nigroviridis, genomic survey sequence.
 ACCESSION AL324737.1 GI:9557619
 VERSION AL324737
 KEYWORDS GSS: genome survey sequence.
 SOURCE Tetraodon nigroviridis
 ORGANISM Tetraodon nigroviridis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodontidae; Tetraodon.

REFERENCE 1
 Roest Crolius, H., Jallion, O., Dasilva, C., Bouneau, L., Fisher, C., Bernot, A., Fizames, C., Wincker, P., Brottier, P., Quetier, F., Saurin, W., and Weissenbach, J.
 Estimate of human gene number provided by genome-wide analysis using Tetraodon nigroviridis DNA sequence
 Nat. Genet. 25 (2), 235-238 (2000)

JOURNAL MEDLINE 20296633
 PUBMED 10835645
 REFERENCE 2
 Roest Crolius, H., Jallion, O., Dasilva, C., Ozouf-Costaz, C., Fizames, C., Fischer, C., Bouneau, L., Billault, A., Quetier, F., Saurin, W., Bernot, A., and Weissenbach, J.
 Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis
 Genome Res. 10 (7), 939-949 (2000)

JOURNAL MEDLINE 20359837
 PUBMED 10899143
 REFERENCE 3 (bases 1 to 1101)
 Genoscope.
 Direct Submission
 Submitted (12-APR-2000) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
 This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at <http://www.genoscope.cns.fr/Tetraodon>.

FEATURES
 .source
 1.1101
 /organism="Tetraodon nigroviridis"
 /mol_type="genomic DNA"
 /db_xref="taxon:99883"
 /clone="002K02"
 /clone_1ib="B"
 /note="Genoscope sequence ID : COAB002BF01BI-end : SP6"

ORIGIN
 Query Match 49.7%; Score 29.8; DB 9; Length 1101;
 Best Local Similarity 70.2%; Pred. No. 39;
 Matches 40; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 3 GTCAACAGCTCTCTCCGCTCATGGGTTACTTGTGCTCCAGATCTGCCAGCTTCC 59
 Db 299 GACTCAGCTCC 243

RESULT 6
 LOCUS BU511928 413 bp mRNA 1linear EST 08-AUG-2002
 DEFINITION BU511928 MF01PSA cDNA Oryzias latipes cDNA clone MF01PSA042E01 3', mRNA sequence.
 ACCESSION BU511928
 VERSION BU511928.1 GI:22163890
 KEYWORDS EST.
 SOURCE Oryzias latipes (Japanese medaka)
 ORGANISM Oryzias latipes
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha; Belontiiformes; Adrianchthyidae; Oryziinae; Oryzias.
 1 (bases 1 to 413)
 Kohara, Y., Shin-I, T., Kimura, T., Narita, T., Uindo, T. and Takeda, H.
 Medaka EST Project in Takeda's lab
 Unpublished (2001)
 Contact: Tadasu Shin-I
 Center For Genetic Resource Information
 National Institute of Genetics
 111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856

Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
1. .413

FEATURES
source
/organism="Oryzias latipes"
/mol_type="mRNA"
/strain="d-r"
/db_xref="taxon:8090"
/clone="MF01FSA042E01"
/sex="mixture of female and male"
/tissue_type="whole embryo"
/dev_stage="fry stage 40"
/clone_1ib="MF01FSA CDNA"

ORIGIN
Query Match 49.3%; Score 29.6; DB 4; Length 413;
Best Local Similarity 68.3%; Pred. No. 42;
Matches 41; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 1 GGGTCACAGCTCTCCCGCTCATGGGTTACTTCTGCTCCAGATACCTCGCCAGCTCCA 60
Db 309 GCGTCATGGGCTCCCGCTCTCATCGGATCTTGGGCTCCAGTCTCGACGAGCTCCA 368

RESULT 7
BU498303/c 568 bp mRNA linear EST 08-AUG-2002
LOCUS BU498303 MF01FSA CDNA Oryzias latipes CDNA clone MF01FSA012104 5',
DEFINITION mRNA sequence.
ACCESSION BU498303
VERSION BU498303.1 GI:22150265
KEYWORDS EST.
SOURCE Oryzias latipes (Japanese medaka)
ORGANISM Oryzias latipes

REFERENCE
AUTHORS Kohara,Y., Shin-I,T., Kimura,T., Narita,T., Jindo,T. and Takeda,H.
TITLES Medaka EST Project in Takeda's lab
JOURNAL Unpublished (2001)
COMMENT Contact: Tadasu Shin-I
Center For Genetic Resource Information
National Institute of Genetics
111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.

FEATURES
source
1. .568
/organism="Oryzias latipes"
/mol_type="mRNA"
/strain="d-r"
/db_xref="taxon:8090"
/clone="MF01FSA012104"
/sex="mixture of female and male"
/tissue_type="whole embryo"
/dev_stage="fry stage 40"
/clone_1ib="MF01FSA CDNA"

ORIGIN

Query Match 49.3%; Score 29.6; DB 4; Length 568;
Best Local Similarity 68.3%; Pred. No. 42;
Matches 41; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 1 GGGTCACAGCTCTCCCGCTCATGGGTTACTTCTGCTCCAGATACCTCGCCAGCTCCA 60
Db 497 GCGTCATGGGCTCCCGCTCTCATCGGATCTTGGGCTCCAGTCTCGACGAGCTCCA 438

RESULT 8
BU500902/c

LOCUS BU500902 584 bp mRNA linear EST 08-AUG-2002
DEFINITION BU500902 MF01FSA CDNA Oryzias latipes CDNA clone MF01FSA047C16 5',
RNA sequence.
ACCESSION BU500902
VERSION BU500902.1 GI:22152864
KEYWORDS EST.
SOURCE Oryzias latipes (Japanese medaka)
ORGANISM Oryzias latipes

REFERENCE
AUTHORS Kohara,Y., Shin-I,T., Kimura,T., Narita,T., Jindo,T. and Takeda,H.
TITLES Medaka EST Project in Takeda's lab
JOURNAL Unpublished (2001)
COMMENT Contact: Tadasu Shin-I
Center For Genetic Resource Information
National Institute of Genetics
111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.

FEATURES
source
1. .584
/organism="Oryzias latipes"
/mol_type="mRNA"
/strain="d-r"
/db_xref="taxon:8090"
/clone="MF01FSA047C16"
/sex="mixture of female and male"
/tissue_type="whole embryo"
/dev_stage="fry stage 40"
/clone_1ib="MF01FSA CDNA"

ORIGIN

Query Match 49.3%; Score 29.6; DB 4; Length 584;
Best Local Similarity 68.3%; Pred. No. 42;
Matches 41; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 1 GGGTCACAGCTCTCCCGCTCATGGGTTACTTCTGCTCCAGATACCTCGCCAGCTCCA 60
Db 524 GCGTCATGGGCTCCCGCTCTCATCGGATCTTGGGCTCCAGTCTCGACGAGCTCCA 465

RESULT 9
BU493595/c 612 bp mRNA linear EST 08-AUG-2002
LOCUS BU493595 MF01FSA CDNA Oryzias latipes CDNA clone MF01FSA032005 5',
DEFINITION mRNA sequence.
ACCESSION BU493595
VERSION BU493595.1 GI:22145521
KEYWORDS EST.
SOURCE Oryzias latipes (Japanese medaka)
ORGANISM Oryzias latipes

REFERENCE
AUTHORS Kohara,Y., Shin-I,T., Kimura,T., Narita,T., Jindo,T. and Takeda,H.
TITLES Medaka EST Project in Takeda's lab
JOURNAL Unpublished (2001)
COMMENT Contact: Tadasu Shin-I
Center For Genetic Resource Information
National Institute of Genetics
111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.

FEATURES
source
1. .612
/organism="Oryzias latipes"

ORIGIN

Query Match 49.3%; Score 29.6; DB 4; Length 612;
 Best Local Similarity 68.3%; Pred. No. 42;
 Matches 41; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

1 GGGTCAACAGCTCTCCCGCTCATGGGCTTACTTGTGCTCCAGATACGCCAGCTCCA 60
 437 GGGTCAACAGCTCTCCCGCTCATGGGCTTACTTGTGCTCCAGATACGCCAGCTCCA 378

/mol_type="mRNA"
 /strain="d-r"
 /db_xref="taxon:8090"
 /clone="MF01FSA032005"
 /sex="mixture of female and male"
 /tissue_type="whole embryo"
 /dev_stage="fry stage 40"
 /clone_lib="MF01FSA CDNA"

ORIGIN

Query Match 49.3%; Score 29.6; DB 4; Length 612;
 Best Local Similarity 68.3%; Pred. No. 42;
 Matches 41; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

1 GGGTCAACAGCTCTCCCGCTCATGGGCTTACTTGTGCTCCAGATACGCCAGCTCCA 60
 437 GGGTCAACAGCTCTCCCGCTCATGGGCTTACTTGTGCTCCAGATACGCCAGCTCCA 378

/mol_type="mRNA"
 /strain="d-r"
 /db_xref="taxon:8090"
 /clone="MF01FSA032005"
 /sex="mixture of female and male"
 /tissue_type="whole embryo"
 /dev_stage="fry stage 40"
 /clone_lib="MF01FSA CDNA"

RESULT 10
 LOCUS BU014441/c 652 bp mRNA linear EST 05-DEC-2001
 DEFINITION BU014441 MF01FSA CDNA Oryzias latipes CDNA clone MF01FSA199H08 5',
 mRNA sequence.
 ACCESSION BU014441
 VERSION BU014441.1 GI:17373347
 KEYWORDS EST
 SOURCE Oryzias latipes (Japanese medaka)
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Oryzias latipes
 Kohara,Y., Shin-I,T., Kimura,T., Narita,T., Jindo,T. and Takeda,H.
 Medaka EST Project in Takeda's lab
 Unpublished (2001)
 Contact: Tadaeu Shin-i
 Center For Genetic Resource Information
 National Institute of Genetics
 111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tshin@genes.nig.ac.jp.

REFERENCE
 AUTHORS Kohara,Y., Shin-I,T., Kimura,T., Narita,T., Jindo,T. and Takeda,H.
 TITLE Medaka EST Project in Takeda's lab
 JOURNAL Unpublished (2001)
 COMMENT Contact: Tadaeu Shin-i
 Center For Genetic Resource Information
 National Institute of Genetics
 111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tshin@genes.nig.ac.jp.

FEATURES
 source
 1..652
 /organism="Oryzias latipes"
 /mol_type="mRNA"
 /strain="d-r"
 /db_xref="taxon:8090"
 /clone="MF01FSA199H08"
 /sex="mixture of female and male"
 /tissue_type="whole embryo"
 /dev_stage="segmentation stage 20 - 25"
 /clone_lib="MF01FSA CDNA"

ORIGIN

Query Match 49.3%; Score 29.6; DB 4; Length 652;
 Best Local Similarity 68.3%; Pred. No. 42;
 Matches 41; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 1 GGGTCAACAGCTCTCCCGCTCATGGGCTTACTTGTGCTCCAGATACGCCAGCTCCA 60
 517 GGGTCAACAGCTCTCCCGCTCATGGGCTTACTTGTGCTCCAGATACGCCAGCTCCA 458

RESULT 11
 LOCUS BU490836/c 652 bp mRNA linear EST 09-AUG-2002
 DEFINITION BU490836 MF01FSA CDNA Oryzias latipes CDNA clone MF01FSA021P19 5',
 mRNA sequence.
 ACCESSION BU490836
 VERSION BU490836.1 GI:22169585

KEYWORDS
 SOURCE EST.
 ORGANISM Oryzias latipes (Japanese medaka)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
 Belontiiformes; Adrianchthyidae; Oryziinae; Oryzias.

REFERENCE
 AUTHORS Kohara,Y., Shin-I,T., Kimura,T., Narita,T., Jindo,T. and Takeda,H.
 TITLE Medaka EST Project in Takeda's lab
 JOURNAL Unpublished (2001)
 COMMENT Contact: Tadaeu Shin-i
 Center For Genetic Resource Information
 National Institute of Genetics
 111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tshin@genes.nig.ac.jp.

FEATURES
 source
 1..652
 /organism="Oryzias latipes"
 /mol_type="mRNA"
 /strain="d-r"
 /db_xref="taxon:8090"
 /clone="MF01FSA021P19"
 /sex="mixture of female and male"
 /tissue_type="whole embryo"
 /dev_stage="fry stage 40"
 /clone_lib="MF01FSA CDNA"

ORIGIN

Query Match 49.3%; Score 29.6; DB 4; Length 652;
 Best Local Similarity 68.3%; Pred. No. 42;
 Matches 41; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 1 GGGTCAACAGCTCTCCCGCTCATGGGCTTACTTGTGCTCCAGATACGCCAGCTCCA 60
 526 GGGTCAACAGCTCTCCCGCTCATGGGCTTACTTGTGCTCCAGATACGCCAGCTCCA 467

RESULT 12
 LOCUS BU504955 669 bp mRNA linear EST 08-AUG-2002
 DEFINITION BU504955 MF01FSA CDNA Oryzias latipes CDNA clone MF01FSA012104 3',
 mRNA sequence.
 ACCESSION BU504955
 VERSION BU504955.1 GI:22156917
 KEYWORDS EST.
 SOURCE Oryzias latipes (Japanese medaka)
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
 Belontiiformes; Adrianchthyidae; Oryziinae; Oryzias.

REFERENCE
 AUTHORS Kohara,Y., Shin-I,T., Kimura,T., Narita,T., Jindo,T. and Takeda,H.
 TITLE Medaka EST Project in Takeda's lab
 JOURNAL Unpublished (2001)
 COMMENT Contact: Tadaeu Shin-i
 Center For Genetic Resource Information
 National Institute of Genetics
 111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tshin@genes.nig.ac.jp.

FEATURES
 source
 1..669
 /organism="Oryzias latipes"
 /mol_type="mRNA"
 /strain="d-r"
 /db_xref="taxon:8090"
 /clone="MF01FSA012104"
 /sex="mixture of female and male"

ORIGIN

/tissue_type="whole embryo"
/dev_stage="fry stage 40"
/clone_id="MF01FSA CDNA"

Query Match 49.3%; Score 29.6; DB 4; Length 669;
Best Local Similarity 68.3%; Pred. No. 43;
Matches 41; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 1 GGGTCAACAGCTCTCCCGCTCATGGGCTTACTTCTGCTCCAGATTAAGTCCGCCAGCTCCA 60
309 GCGTCATGGGCTCCCGCTCTCATCGGATCTTCGGGCTCCAGATGCTCGACGAGCTCCA 368

RESULT 13

LOCUS B028611 686 bp mRNA linear EST 06-DEC-2001
DEFINITION B028611 MF01FSA CDNA Oryzias latipes CDNA clone MF01SSA19H08 3',
mRNA sequence.

ACCESSION B028611
VERSION B028611
KEYWORDS GI:17396355
SOURCE EST

ORGANISM Oryzias latipes (Japanese medaka)

REFERENCE Oryzias latipes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percormorpha; Atherinomorpha;
Belontiiformes; Adrianiichthyidae; Oryziinae; Oryzias.

1 (bases 1 to 686)
Kohara,Y., Shin-I,T., Kimura,T., Narita,T., Jindo,T. and Takeda,H.
Medaka EST Project in Takeda's lab
Unpublished (2001)
Contact: Tadasu Shin-I

Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.

FEATURES

source

1..686
/organism="Oryzias latipes"
/mol_type="mRNA"
/strain="Hd-rR"
/db_xref="taxon:8090"
/clone="MF01SSA19H08"
/sex="mixture of female and male"
/tissue_type="whole embryo"
/dev_stage="segmentation stage 20 - 25"
/clone_id="MF01FSA CDNA"

ORIGIN

Query Match 49.3%; Score 29.6; DB 4; Length 686;
Best Local Similarity 68.3%; Pred. No. 43;
Matches 41; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 1 GGGTCAACAGCTCTCCCGCTCATGGGCTTACTTCTGCTCCAGATTAAGTCCGCCAGCTCCA 60
324 GCGTCATGGGCTCCCGCTCTCATCGGATCTTCGGGCTCCAGATGCTCGACGAGCTCCA 383

RESULT 14

LOCUS B0710454 697 bp mRNA linear EST 08-MAR-2004
DEFINITION B0710454 MF01FPA CDNA Oryzias latipes CDNA clone MF01FPA023111 5',
mRNA sequence.

ACCESSION B0710454
VERSION B0710454
KEYWORDS GI:45251398
SOURCE EST

ORGANISM Oryzias latipes (Japanese medaka)

REFERENCE Oryzias latipes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

REFERENCE Acanthomorpha; Acanthopterygii; Percormorpha; Atherinomorpha;
Belontiiformes; Adrianiichthyidae; Oryziinae; Oryzias.
1 (bases 1 to 697)
Kohara,Y., Shin-I,T., Kimura,T., Narita,T., Jindo,T. and Takeda,H.
Medaka EST Project in Takeda's lab
Unpublished (2001)
Contact: Tadasu Shin-I

Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.

FEATURES

source

1..697
/organism="Oryzias latipes"
/mol_type="mRNA"
/strain="Hd-rR"
/db_xref="taxon:8090"
/clone="MF01FPA023111"
/sex="mixture of female and male"
/tissue_type="whole embryo"
/dev_stage="fry stage 40"
/clone_id="MF01FPA CDNA"

ORIGIN

Query Match 49.3%; Score 29.6; DB 4; Length 697;
Best Local Similarity 68.3%; Pred. No. 43;
Matches 41; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 1 GGGTCAACAGCTCTCCCGCTCATGGGCTTACTTCTGCTCCAGATTAAGTCCGCCAGCTCCA 60
552 GCGTCATGGGCTCCCGCTCTCATCGGATCTTCGGGCTCCAGATGCTCGACGAGCTCCA 493

RESULT 15

LOCUS B0508359 730 bp mRNA linear EST 08-AUG-2002
DEFINITION B0508359 MF01FSA CDNA Oryzias latipes CDNA clone MF01FSA032005 3',
mRNA sequence.

ACCESSION B0508359
VERSION B0508359
KEYWORDS GI:22160321
SOURCE EST

ORGANISM Oryzias latipes (Japanese medaka)

REFERENCE Oryzias latipes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percormorpha; Atherinomorpha;
Belontiiformes; Adrianiichthyidae; Oryziinae; Oryzias.

1 (bases 1 to 730)
Kohara,Y., Shin-I,T., Kimura,T., Narita,T., Jindo,T. and Takeda,H.
Medaka EST Project in Takeda's lab
Unpublished (2001)
Contact: Tadasu Shin-I

Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.

FEATURES

source

1..730
/organism="Oryzias latipes"
/mol_type="mRNA"
/strain="d-rR"
/db_xref="taxon:8090"
/clone="MF01FSA032005"
/sex="mixture of female and male"
/tissue_type="whole embryo"
/dev_stage="fry stage 40"
/clone_id="MF01FSA CDNA"

ORIGIN

This Page Blank (uspto)

XX
SQ Sequence 60 BP; 10 A; 23 C; 13 G; 14 T; 0 U; 0 Other;
XX
Query Match 100.0%; Score 60; DB 2; Length 60;
XX Best Local Similarity 100.0%; Pred. No. 3.7e-10;
XX Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 GGGTCAACAGCTCTCCGCTCATGGGTTACTTCTGCTCCAGATATCTGGCCAGCTCCA 60
DB 1 GGGTCAACAGCTCTCCGCTCATGGGTTACTTCTGCTCCAGATATCTGGCCAGCTCCA 60
RESULT 2
ID ADK68635/C
ADK68635 standard; cDNA; 156 BP.
XX
AC ADK68635;
XX
DT 06-MAY-2004 (first entry)
XX
DE HSP65-MUCL antigen CTL epitope related cDNA #3.
XX
KM Human breast cancer; Mycobacterium bovis; heat shock protein 65; BCG;
KW HSP65; cytotoxic T lymphocyte; CTL; MUCL; HSP65-MUCL antigen CTL epitope;
XX ss.
XX
OS Unidentified.
XX
PN CN1368384-A.
XX
PD 11-SEP-2002.
XX
PF 08-FEB-2001; 2001CN-00102614.
XX
PR 08-FEB-2001; 2001CN-00102614.
XX
PA (DIWE-) DIWEIHUAYU BIO TECHNOLOGY CO LTD BEIJING.
XX
PI Yu Y, Li H;
XX
DR WPI; 2003-854662/80.
XX
PT Genetically engineered vaccine of MUC-1 antigen for human breast cancer.
XX
PS Example 2; Page 4 (Disclosure); 14pp; Chinese.
XX
CC The invention relates to a method of preparation of a genetically
CC engineered vaccine for preventing and treating human breast cancer. The
CC method comprises fusing the coding gene of the Mycobacterium bovis heat
CC shock protein 65 (BCG HSP65) with the cytotoxic T lymphocyte (CTL)
CC epitope gene of MUCL (antigen cell expressed by human breast cancer
CC cells) and the HSP65-MUCL antigen CTL epitope is expressed in Escherichia
CC coli cells. This sequence represents DNA used in the method of the
CC invention.
XX
SQ Sequence 156 BP; 37 A; 42 C; 58 G; 19 T; 0 U; 0 Other;
XX
Query Match 63.7%; Score 38.2; DB 10; Length 156;
XX Best Local Similarity 78.0%; Pred. No. 0.0059;
XX Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
OY 1 GGGTCAACAGCTCTCCGCTCATGGGTTACTTCTGCTCCAGATATCTGGCCAGCTCC 59
DB 132 GGTCTACCGCTCCGCGCTCAGCGTGTACTCTGCTCCGACACCGCTCCGCTCC 74
RESULT 3
ID ADK68629
ADK68629 standard; cDNA; 1800 BP.
XX
AC ADK68629;
XX
DT 06-MAY-2004 (first entry)

XX
DE HSP65-MUCL antigen CTL epitope related cDNA #1.
XX
KM Human breast cancer; Mycobacterium bovis; heat shock protein 65; BCG;
KW HSP65; cytotoxic T lymphocyte; CTL; MUCL; HSP65-MUCL antigen CTL epitope;
XX gene; ss.
XX
OS Unidentified.
XX
PN CN1368384-A.
XX
PD 11-SEP-2002.
XX
PF 08-FEB-2001; 2001CN-00102614.
XX
PR 08-FEB-2001; 2001CN-00102614.
XX
PA (DIWE-) DIWEIHUAYU BIO TECHNOLOGY CO LTD BEIJING.
XX
PI Yu Y, Li H;
XX
DR WPI; 2003-854662/80.
XX
DR P-PsDB; ADK68630.
XX
PT Genetically engineered vaccine of MUC-1 antigen for human breast cancer.
XX
PS Claim 1; SEQ ID NO 1; 14pp; Chinese.
XX
CC The invention relates to a method of preparation of a genetically
CC engineered vaccine for preventing and treating human breast cancer. The
CC method comprises fusing the coding gene of the Mycobacterium bovis heat
CC shock protein 65 (BCG HSP65) with the cytotoxic T lymphocyte (CTL)
CC epitope gene of MUCL (antigen cell expressed by human breast cancer
CC cells) and the HSP65-MUCL antigen CTL epitope is expressed in Escherichia
CC coli cells. This sequence represents DNA used in the method of the
CC invention.
XX
SQ Sequence 1800 BP; 347 A; 561 C; 616 G; 276 T; 0 U; 0 Other;
XX
Query Match 63.7%; Score 38.2; DB 10; Length 1800;
XX Best Local Similarity 78.0%; Pred. No. 0.0088;
XX Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
OY 1 GGGTCAACAGCTCTCCGCTCATGGGTTACTTCTGCTCCAGATATCTGGCCAGCTCC 59
DB 1639 GGTCTACCGCTCCGCGCTCAGGTTACTTCTGCTCCGACACCGCTCCGCTCC 1697
RESULT 4
ID AAV48324
AAV48324 standard; DNA; 60 BP.
XX
AC AAV48324;
XX
DT 20-NOV-1998 (first entry)
XX
DE Nucleotide sequence encoding MUCL tandem repeat unit R9.
XX
KW ss; MUCL; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
KW tumour; tumour-associated antigen.
XX
OS Homo sapiens.
XX
PN W09837095-A2.
XX
PD 27-AUG-1998.
XX
PF 24-FEB-1998; 98MO-US003693.
XX
PR 24-FEB-1997; 97US-0038253P.
XX
PA (THER-) THERION BIOLOGICS CORP.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

PA (DAND) DANA FARBER CANCER INST INC.
 XX
 PI Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;
 XX
 DR WPI; 1998-467492/40.
 XX
 PT New recombinant pox virus for tumour therapy - comprises DNA encoding an
 PT immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
 XX
 PS Disclosure; Page 11; 42pp; English.
 XX
 CC The MUC1 tandem repeat units AAV48317-V48325 were used to create an
 CC immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
 CC (RPV). The RPV was used in a pharmaceutical composition also containing
 CC an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The
 CC recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
 CC does not undergo significant genetic deletion, thereby providing an
 CC unexpectedly stable and immunogenic pox virus. They can be used to
 CC prevent or treat tumours expressing MUC1 tumour-associated antigens
 XX
 SQ Sequence 60 BP; 12 A; 26 C; 14 G; 8 T; 0 U; 0 Other;
 XX
 Query Match 61.0%; Score 36.6; DB 2; Length 60;
 Best Local Similarity 76.3%; Pred. No. 0.017;
 Matches 45; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
 XX
 OY 1 GGGTCAACAGCTCTCCGCTCATGGGTTACTTCTGCTCCAGATACGCCAGCTCC 59
 1 GGTTCACGCGACCTCCAGCAGCAGGATGACGCTGACACCGACCCGCTCAGCTCC 59
 DB
 RESULT 5
 AAV48323
 ID AAV48323 standard; DNA; 60 BP.
 XX
 AC AAV48323;
 XX
 DT 20-NOV-1998 (first entry)
 XX
 DE Nucleotide sequence encoding MUC1 tandem repeat unit R8.
 XX
 XX ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
 XX tumour; tumour-associated antigen.
 XX
 OS Homo sapiens.
 XX
 PN WO9837095-A2.
 XX
 PD 27-AUG-1998.
 XX
 PF 24-FEB-1998; 98WO-US003693.
 XX
 PR 24-FEB-1997; 97US-0038253P.
 XX
 PA (THER-) THERION BIOLOGICS CORP.
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA (DAND) DANA FARBER CANCER INST INC.
 XX
 PI Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;
 XX
 DR WPI; 1998-467492/40.
 XX
 PT New recombinant pox virus for tumour therapy - comprises DNA encoding an
 PT immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
 XX
 PS Disclosure; Page 11; 42pp; English.
 XX
 CC The MUC1 tandem repeat units AAV48317-V48325 were used to create an
 CC immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
 CC (RPV). The RPV was used in a pharmaceutical composition also containing
 CC an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The
 CC recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
 CC does not undergo significant genetic deletion, thereby providing an

CC unexpectedly stable and immunogenic pox virus. They can be used to
 CC prevent or treat tumours expressing MUC1 tumour-associated antigens
 XX
 SQ Sequence 60 BP; 10 A; 24 C; 17 G; 9 T; 0 U; 0 Other;
 XX
 Query Match 60.0%; Score 36; DB 2; Length 60;
 Best Local Similarity 75.0%; Pred. No. 0.026;
 Matches 45; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
 XX
 OY 1 GGGTCAACAGCTCTCCGCTCATGGGTTACTTCTGCTCCAGATACGCCAGCTCC 60
 1 GGTTCAGACTGCGCTCCGCGCATGATGTGACCTCAGCTCCGACACAGCGCCGCCA 60
 DB
 RESULT 6
 ADK68631
 ID ADK68631 standard; cDNA; 120 BP.
 XX
 AC ADK68631;
 XX
 DT 06-MAY-2004 (first entry)
 XX
 DE HSP65-MUC1 antigen CTL epitope related cDNA #2.
 XX
 KW Human breast cancer; Mycobacterium bovis; heat shock protein 65; BCG;
 KW HSP65; cytotoxic T lymphocyte; CTL; MUC1; HSP65-MUC1 antigen CTL epitope;
 KW ss.
 XX
 OS unidentified.
 XX
 PN CN1368384-A.
 XX
 PD 11-SEP-2002.
 XX
 PF 08-FEB-2001; 2001CN-00102614.
 XX
 PR 08-FEB-2001; 2001CN-00102614.
 XX
 PA (DIWE-) DIWEIHUAYU BIO TECHNOLOGY CO LTD BEIJING.
 XX
 PI Yu Y, Li H;
 XX
 DR WPI; 2003-854662/80.
 XX
 PT Genetically engineered vaccine of MUC-1 antigen for human breast cancer.
 XX
 PS Claim 4; SEQ ID NO 3; 14pp; Chinese.
 XX
 CC The invention relates to a method of preparation of a genetically
 CC engineered vaccine for preventing and treating human breast cancer. The
 CC method comprises fusing the coding gene of the Mycobacterium bovis heat
 CC shock protein 65 (BCG HSP65) with the cytotoxic T lymphocyte (CTL)
 CC epitope gene of MUC1 (antigen cell expressed by human breast cancer
 CC cells) and the HSP65-MUC1 antigen CTL epitope is expressed in Escherichia
 CC coli cells. This sequence represents DNA used in the method of the
 CC invention.
 XX
 SQ Sequence 120 BP; 11 A; 43 C; 37 G; 27 T; 0 U; 2 Other;
 XX
 Query Match 56.0%; Score 33.6; DB 10; Length 120;
 Best Local Similarity 72.4%; Pred. No. 0.18;
 Matches 42; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
 XX
 OY 1 GGGTCAACAGCTCTCCGCTCATGGGTTACTTCTGCTCCAGATACGCCAGCTC 58
 61 GGTTCACGCGCTCCGCGCATGATGTGACCTCAGCTCCGACACAGCGCTCCGCTC 118
 DB
 RESULT 7
 AAV48318
 ID AAV48318 standard; DNA; 60 BP.
 XX
 AC AAV48318;

```
XX 20-NOV-1998 (first entry)
DT Nucleotide sequence encoding MUC1 tandem repeat unit R3.
DE MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
XX 56; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
XX tumour; tumour-associated antigen.
XX Homo sapiens.
OS WO9837095-A2.
XX 27-AUG-1998.
XX 24-FEB-1998; 98WO-US003693.
XX 24-FEB-1997; 97US-0038253P.
XX (THER-) THERION BIOLOGICS CORP.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX (DAND) DANA FARBER CANCER INST INC.
XX Schlom J, Kantar J, Kufe D, Panicali D, Gritz L;
XX WPI; 1998-467492/40.
XX New recombinant pox virus for tumour therapy - comprises DNA encoding an
XX immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
XX Disclosure; Page 11; 42pp; English.
XX The MUC1 tandem repeat units AAV48317-V48325 were used to create an
XX immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
XX (RPV). The RPV was used in a pharmaceutical composition also containing
XX an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The
XX recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
XX does not undergo significant genetic deletion, thereby providing an
XX unexpectedly stable and immunogenic pox virus. They can be used to
XX prevent or treat tumours expressing MUC1 tumour-associated antigens
XX
SQ Sequence 60 BP; 7 A; 29 C; 19 G; 5 T; 0 U; 0 Other;
Query Match 53.0%; Score 31.8; DB 2; Length 60;
Best Local Similarity 71.2%; Pred. No. 0.62; Mismatches 0; Gaps 0;
Matches 42; Conservative 0; Indels 17; Indels 0; Gaps 0;
QY 1 GGGTCAACAGCTCTCCGCTCATGGGTTACTTCTGCTCAGATACCTGCCAGCTCC 59
DB 1 GGATCCACCGCGCCGCTCGCACGAGTGAACGTGGCGCCCGACACCGCGCCGCTCC 59
RESULT 8
AAD00385 ID AAD00385 standard; DNA; 525 BP.
XX
AC AAD00385;
XX
DT 29-AUG-2000 (first entry)
XX
DE Human Mucin 1 (MUC-1) protein fragment encoding DNA #2.
XX
KW Human; Mucin 1; MUC-1; tumour; pMRS30 expression vector; anti-tumour;
XX therapy; immune response; cytostatic; vaccine; ds.
XX Homo sapiens.
OS
XX
FH Key Location/Qualifiers
FT 1..525
FT /tag= a
FT /product= "MUC-1 protein fragment"
XX
XX WO200025827-A2.
```

```
PD 11-MAY-2000.
XX 18-OCT-1999; 99WO-EP007874.
XX 30-OCT-1998; 98IT-MI002330.
XX (MENA) MENARINI RICERCHE SPA.
XX Parente D, Di Massimo AM, De Santis R;
XX WPI; 2000-365410/31.
XX P-PSDB; AAV71021.
XX Composition containing one or more DNA molecules encoding fragments of a
XX Mucin 1 (MUC-1) protein overexpressed in tumor cells, useful in anti-
XX tumor therapy.
XX Claim 16; Fig 2; 56pp; English.
XX The present sequence is a DNA encoding a fragment of human Mucin 1 (MUC-
XX 1) antigenic protein which is overexpressed in tumour cells. The sequence
XX was obtained from ER20 tumour cells by reverse transcriptase-PCR and
XX corresponds to nucleotides 205-720 of the EMBL sequence J05581 with a
XX start codon and two stop codons. The present sequence is cloned into a
XX pMRS30 expression vector and used in pharmaceutical composition e.g.
XX vaccine for inducing an antigen-specific anti-tumour immune response.
XX Composition containing this DNA molecule is useful in anti-tumour therapy
XX of patients affected with tumours characterised by high MUC-1 expression
XX
SQ Sequence 525 BP; 111 A; 199 C; 128 G; 87 T; 0 U; 0 Other;
Query Match 52.7%; Score 31.6; DB 3; Length 525;
Best Local Similarity 74.1%; Pred. No. 1;
Matches 40; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 7 ACAGCTCTCCGCGCTCATGGGTTACTTCTGCTCAGATACCTGCCAGCTCCA 60
DB 262 ACTGCTCACGACGACACGAGTGTACTCTCGGCTCGGATACGAGCGGCCCA 315
RESULT 9
AAD00391 ID AAD00391 standard; DNA; 891 BP.
XX
AC AAD00391;
XX
DT 15-SEP-2003 (revised)
DT 29-AUG-2000 (first entry)
XX
DE Ubiquitin-E. coli LacI-human Mucin 1 fusion protein encoding DNA #2.
XX
KW Ubiquitin; LacI; beta-galactosidase; fusion protein; human; Mucin 1;
XX MUC-1; tumour; pMRS30 expression vector; anti-tumour; therapy;
XX immune response; cytostatic; vaccine; ds.
XX Homo sapiens.
XX Escherichia coli.
XX Chimeric.
XX
FH Key Location/Qualifiers
FT 1..891
FT /tag= a
FT /product= "Ubiquitin-LacI-MUC-1 fusion protein"
FT 1..369
FT /tag= b
FT /label= ubiLacI DNA
FT /note= "Includes ubiquitin-E. coli LacI fusion DNA"
FT 370..891
FT /tag= c
FT /note= "Human MUC-1 partial DNA that corresponds to
XX nucleotides 205-720 of the EMBL sequence J05581 with two
XX stop codons"
```

PN WO200025827-A2.
 XX
 PD 11-MAY-2000.
 XX
 PF 18-OCT-1999; 99WO-EP007874.
 XX
 PR 30-OCT-1998; 98IT-MI002330.
 XX
 PA (MENA) MENARINI RICERCHE SPA.
 XX
 PI Parente D, Di Massimo AM, De Santis R;
 XX
 DR WPI: 2000-365410/31.
 DR P-PSDB; AAY71027.
 XX
 PT Composition containing one or more DNA molecules encoding fragments of a
 PT Mucin 1 (MUC-1) protein overexpressed in tumor cells, useful in anti-
 PT tumor therapy.
 XX
 PS Claim 18; Fig 8; 56pp; English.
 XX
 CC The present sequence is a DNA encoding a fusion protein consisting of
 CC human Mucin 1 (MUC-1) fragment fused to Ubilacti sequence at the N-
 CC terminus. The Ubilacti sequence consists of ubiquitin from MCF7 cell line
 CC and a portion of E. coli beta-galactosidase (LacI). MUC-1 is an antigenic
 CC protein overexpressed in tumor cells. The present sequence is cloned
 CC into a pMR30 expression vector and used in pharmaceutical composition
 CC e.g. vaccine for inducing an antigen-specific anti-tumour immune
 CC response. Composition containing this DNA molecule is useful in anti-
 CC tumour therapy of patients affected with tumours characterised by high
 CC MUC-1 expression. (Updated on 15-SEP-2003 to standardise OS field)
 CC
 SQ Sequence 891 BP; 203 A; 307 C; 225 G; 156 T; 0 U; 0 Other;
 SQ
 Query Match 52.7%; Score 31.6; DB 3; Length 891;
 Best Local Similarity 74.1%; Pred. No. 1.1;
 Matches 40; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
 QY 7 ACAGCTCTCCCGCTCATGGGTTACTTCTGCTCCAGATCTCGCCAGCTCCA 60
 DB 628 ACTGCTCCACGACGACGCGTGTACTCTCGCTCCGATATCCAGCGCGGCCCA 681

RESULT 10
 AAD00388
 ID AAD00388 standard; DNA; 1371 BP.
 AC
 XX AAD00388;
 XX
 DT 29-AUG-2000 (first entry)
 XX
 DE Human Mucin 1 (MUC-1) protein fragment encoding DNA #5.
 XX
 KW Human; Mucin 1; MUC-1; tumour; pMR30 expression vector; anti-tumour;
 KW therapy; immune response; cytostatic; vaccine; ds.
 XX
 OS Homo sapiens.
 OS
 FH Key Location/Qualifiers
 FT CDS 1..1371
 FT /tag= a
 FT /product= "MUC-1 protein fragment"
 FT
 PN WO200025827-A2.
 PD 11-MAY-2000.
 XX
 PF 18-OCT-1999; 99WO-EP007874.
 XX
 PR 30-OCT-1998; 98IT-MI002330.
 XX
 PA (MENA) MENARINI RICERCHE SPA.
 XX

PI Parente D, Di Massimo AM, De Santis R;
 XX
 DR WPI: 2000-365410/31.
 DR P-PSDB; AAY71024.
 XX
 PT Composition containing one or more DNA molecules encoding fragments of a
 PT Mucin 1 (MUC-1) protein overexpressed in tumor cells, useful in anti-
 PT tumor therapy.
 XX
 PS Claim 16; Fig 5; 56pp; English.
 XX
 CC The present sequence is a DNA encoding a fragment of human Mucin 1 (MUC-
 CC 1) antigenic protein which is overexpressed in tumor cells. The sequence
 CC was obtained by PCR from plasmids pMR3166, pMR3167, pMR3168 and pMR3169
 CC which contain MUC-1 DNA from BT20 tumour cells. It corresponds to
 CC nucleotides 136-1497 of the EMBL sequence J05581 with a start codon and
 CC two stop codons. The present sequence is cloned into a pMR30 expression
 CC vector and used in pharmaceutical composition e.g. vaccine for inducing
 CC an antigen-specific anti-tumour immune response. Composition containing
 CC this DNA molecule is useful in anti-tumour therapy of patients affected
 CC with tumours characterised by high MUC-1 expression
 CC
 SQ Sequence 1371 BP; 300 A; 450 C; 321 G; 300 T; 0 U; 0 Other;
 SQ
 Query Match 52.7%; Score 31.6; DB 3; Length 1371;
 Best Local Similarity 74.1%; Pred. No. 1.2;
 Matches 40; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
 QY 7 ACAGCTCTCCCGCTCATGGGTTACTTCTGCTCCAGATCTCGCCAGCTCCA 60
 DB 331 ACCGCTCCACGACGACGCGTGTACTCTCGCTCCGATATCCAGCGCGGCCCA 384

RESULT 11
 AAD00394
 ID AAD00394 standard; DNA; 1737 BP.
 AC
 XX AAD00394;
 XX
 DT 15-SEP-2003 (revised)
 DT 29-AUG-2000 (first entry)
 XX
 DE Ubiquitin-E. coli LacI-human Mucin 1 fusion protein encoding DNA #5.
 XX
 KW Ubiquitin; LacI; beta-galactosidase; fusion protein; human; Mucin 1;
 KW MUC-1; tumour; pMR30 expression vector; anti-tumour; therapy;
 KW immune response; cytostatic; vaccine; ds.
 XX
 OS Homo sapiens.
 OS
 OS Escherichia coli.
 OS Chimeric.
 OS
 FH Key Location/Qualifiers
 FT CDS 1..1737
 FT /tag= a
 FT /product= "Ubiquitin-LacI-MUC-1 fusion protein"
 FT
 FT misc_feature 1..369
 FT /tag= b
 FT /label= UBILACTI DNA
 FT /note= "Includes ubiquitin-E. coli LacI fusion DNA"
 FT
 FT misc_feature 370..1737
 FT /tag= c
 FT /note= "Human MUC-1 partial DNA that corresponds to
 FT nucleotides 136-1497 of the EMBL sequence J05581 with two
 FT stop codons"
 FT
 PN WO200025827-A2.
 PD 11-MAY-2000.
 XX
 PF 18-OCT-1999; 99WO-EP007874.
 XX
 PR 30-OCT-1998; 98IT-MI002330.
 XX

XX (MENA) MENARINI RICERCH SPA.
 XX Parente D, Di Massimo AM, De Santis R;
 FI WPI; 2000-365410/31.
 DR P-PSDB; AAV71030.
 XX Composition containing one or more DNA molecules encoding fragments of a
 PT Mucin 1 (MUC-1) protein overexpressed in tumor cells, useful in anti-
 PT tumor therapy.
 PS Claim 18; Fig 11; 56pp; English.
 XX The present sequence is a DNA encoding a fusion protein consisting of
 CC human Mucin 1 (MUC-1) fragment fused to UBIlacti sequence at the N-
 CC terminus. The UBIlacti sequence consists of ubiquitin from MCF7 cell line
 CC and a portion of E. coli beta-galactosidase (LactI). MUC-1 is an antigenic
 CC protein overexpressed in tumour cells. The present sequence is cloned
 CC into a pMR30 expression vector and used in pharmaceutical composition
 CC e.g. vaccine for inducing an antigen-specific anti-tumour immune
 CC response. Composition containing this DNA molecule is useful in anti-
 CC tumour therapy of patients affected with tumours characterised by high
 CC MUC-1 expression. (Updated on 15-SEP-2003 to standardise OS field)
 XX Sequence 1737 BP; 392 A; 558 C; 418 G; 369 T; 0 U; 0 Other;
 SQ
 Query Match 52.7%; Score 31.6; DB 3; Length 1737;
 Best Local Similarity 74.1%; Pred. No. 1.3; 14; Indels 0; Gaps 0;
 Matches 40; Conservative 0; Mismatches 0;
 Oy 7 AAGAGCTCTCCGCTCATGGGGTTACTTGTCTCCAGATACCTGCCAGCTCCA 60
 Db 697 ACCGCTCCACGACGACGAGGTGTACCTCGCTCCGATACGAGCGGCCCA 750
 RESULT 12
 AAV48320
 ID AAV48320 standard; DNA; 60 BP.
 AC AAV48320;
 DT 20-NOV-1998 (first entry)
 XX Nucleotide sequence encoding MUC1 tandem repeat unit R5.
 DB ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
 KM tumour; tumour-associated antigen.
 XX Homo sapiens.
 OS
 XX WO9837095-A2.
 PN 27-AUG-1998.
 PD 24-FEB-1998; 98WO-US003693.
 PF 24-FEB-1997; 97US-0038253P.
 PR 24-FEB-1997;
 XX (THER-) THERION BIOLOGICS CORP.
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA (DAND) DANA FARBBER CANCER INST INC.
 XX Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;
 PI WPI; 1998-467492/40.
 DR New recombinant pox virus for tumour therapy - comprises DNA encoding an
 PT immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
 XX Disclosure; Page 11; 42pp; English.
 CC The MUC1 tandem repeat units AAV48317-V48325 were used to create an

CC immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
 CC (RPV). The RPV was used in a pharmaceutical composition also containing
 CC an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The
 CC recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
 CC does not undergo significant genetic deletion, thereby providing an
 CC unexpectedly stable and immunogenic pox virus. They can be used to
 CC prevent or treat tumours expressing MUC1 tumour-associated antigens
 XX Sequence 60 BP; 9 A; 27 C; 15 G; 9 T; 0 U; 0 Other;
 SQ
 Query Match 50.3%; Score 30.2; DB 2; Length 60;
 Best Local Similarity 69.5%; Pred. No. 2.1;
 Matches 41; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
 Oy 1 GGGTCAACAGCTCCTCCGCTCATGGGGTTACTTGTCTCCAGATACCTGCCAGCTCC 59
 Db 1 GGTTCGACGGGCCCCCTGCTCAGGAGTGTACATCCGCCGAGTACAGACGCGCCCC 59
 RESULT 13
 ID AAS01385 standard; DNA; 68 BP.
 AC AAS01385;
 DT 04-JUL-2001 (first entry)
 XX SBMV coat protein-MUC(16) insertion construct #3.
 DE Human; polymorphic epithelial cell mucin; PEM; chimeric virus particle;
 XX CVP; plant virus coat protein; comovirus; CPNV; cowpea mosaic virus;
 KM SBMV; Southern bean mosaic virus; LRSV; red clover necrotic mosaic virus;
 KM RCNMV; MUC1; tumour; cancer vaccine; mutant; ds.
 XX Southern bean mosaic virus.
 OS Homo sapiens.
 OS Synthetic.
 OS Chimeric.
 OS
 FH Key Location/Qualifiers
 FT misc_recomb 12..59
 FT /*tag= a
 FT /note= "Human MUC(16) sequence (AAS01380)"
 XX WO200118199-A1.
 PN 15-MAR-2001.
 PD 11-SEP-2000; 2000WO-GB003500.
 PF 09-SEP-1999; 99GB-00021337.
 PR (DOWC) DOW CHEM CO.
 PA Bendig MM, Jones TD, Longstaff M, Hellebrand K;
 PI WPI; 2001-244570/25.
 DR New chimeric plant virus particles with an immunogenically active peptide
 PT of a tumor-associated mucin, useful as a vaccine or for the manufacture
 PT of a vaccine for treating and/or preventing tumors and/or cancer.
 XX Example 12; Page; 63pp; English.
 PS The present sequence for SBMV (Southern bean mosaic virus) coat protein-
 CC MUC(16) insertion construct #3 is 1 of 5 constructs containing a MUC1(16)
 CC DNA sequence at different positions within a SBMV coat protein insertion
 CC sequence (AAS01382). The construct is used to create a novel chimeric
 CC virus particle (CVP). MUC1(16) peptide is 1 of 7 polymorphic epithelial
 CC cell mucin (PEM) peptides (AAU00483-AAU00489) which can be used to create
 CC CVPs. The mucin peptide epitopes are preferably inserted into the coat
 CC protein of a plant virus such as the comovirus CPNV (cowpea mosaic
 CC virus). The ability of the chimeric virus particle CPNV-MUC1(16) to

CC elicit antibodies, which can cause regression of tumours expressing the
 CC MUC1 protein, is demonstrated in a mouse tumour model. Other examples of
 CC CVPs include the insertion of the mucin MUC1(16) peptide into the coat
 CC protein of other plant viruses e.g. LRSV (AAU00491) and RCNMV (red clover
 CC necrotic mosaic virus; AAU00492). The CVP is useful as or as part of a
 CC vaccine particularly for treating and preventing tumours and cancer. The
 CC CVP provides advantages over prior art antigen-presenting means since
 CC conventional live animal virus vectors can be avoided; as can the need
 CC for separate mucin peptide synthesis and chemical-coupling to a
 CC conventional carrier. Also, the CVP is shown to induce good mucosal
 CC immunity. Note: The present sequence is not given in the patent but is
 CC indexed from the information provided

XX
 SQ Sequence 68 BP; 10 A; 19 C; 14 G; 25 T; 0 U; 0 Other;

Query Match 50.3%; Score 30.2; DB 4; Length 68;
 Best Local Similarity 81.4%; Pred. No. 2.1;
 Matches 35; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 17 CCGCTCATGGGGTCTCTGCTCCAGATCTGCCCACTCC 59
 Db 4 CCTTAAGGGTGTACTCTGCTCCGATCTAGACTCTCTCC 46

RESULT 14

ID AAV48321 standard; DNA; 60 BP.

XX AAV48321;

XX 20-NOV-1998 (first entry)

XX Nucleotide sequence encoding MUC1 tandem repeat unit R6.

XX 86; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
 XX tumour; tumour-associated antigen.

XX Homo sapiens.

XX WO9837095-A2.

XX 27-AUG-1998.

XX 24-FEB-1998; 98WO-US003693.

XX 24-FEB-1997; 97US-0038253P.

XX (THER-) THERION BIOLOGICS CORP.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX (DAND) DANA FARBER CANCER INST INC.

XX Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;

XX WPI; 1998-467492/40.

XX New recombinant pox virus for tumour therapy - comprises DNA encoding an
 XX immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.

XX Disclosure; Page 11; 42pp; English.

XX The MUC1 tandem repeat units AAV48317-V48325 were used to create an
 CC immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
 CC (RPV). The RPV was used in a pharmaceutical composition also containing
 CC an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The
 CC recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
 CC does not undergo significant genetic deletion, thereby providing an
 CC unexpectedly stable and immunogenic pox virus. They can be used to
 CC prevent or treat tumours expressing MUC1 tumour-associated antigens

XX Sequence 60 BP; 13 A; 27 C; 17 G; 3 T; 0 U; 0 Other;

Query Match 50.0%; Score 30; DB 2; Length 60;
 Best Local Similarity 72.2%; Pred. No. 2.4;

Matches 39; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 7 ACAGCTCTCCCGCTCATGGGGTGTACTCTGCTCCAGATCTGCCCACTCC 60
 Db 7 ACCGCTCATCCCGCTCATGGGGTGTACTCTGCTCCAGATCTGCCCACTCC 60

RESULT 15

ID AAV48322 standard; DNA; 60 BP.

XX AAV48322;

XX 20-NOV-1998 (first entry)

XX Nucleotide sequence encoding MUC1 tandem repeat unit R7.

XX 86; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
 XX tumour; tumour-associated antigen.

XX Homo sapiens.

XX WO9837095-A2.

XX 27-AUG-1998.

XX 24-FEB-1998; 98WO-US003693.

XX 24-FEB-1997; 97US-0038253P.

XX (THER-) THERION BIOLOGICS CORP.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX (DAND) DANA FARBER CANCER INST INC.

XX Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;

XX WPI; 1998-467492/40.

XX New recombinant pox virus for tumour therapy - comprises DNA encoding an
 XX immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.

XX Disclosure; Page 11; 42pp; English.

XX The MUC1 tandem repeat units AAV48317-V48325 were used to create an
 CC immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
 CC (RPV). The RPV was used in a pharmaceutical composition also containing
 CC an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The
 CC recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
 CC does not undergo significant genetic deletion, thereby providing an
 CC unexpectedly stable and immunogenic pox virus. They can be used to
 CC prevent or treat tumours expressing MUC1 tumour-associated antigens

XX Sequence 60 BP; 14 A; 24 C; 16 G; 6 T; 0 U; 0 Other;

Query Match 50.0%; Score 30; DB 2; Length 60;
 Best Local Similarity 72.2%; Pred. No. 2.4;
 Matches 39; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 7 ACAGCTCTCCCGCTCATGGGGTGTACTCTGCTCCAGATCTGCCCACTCC 60
 Db 7 ACCGCTCATCCCGCTCATGGGGTGTACTCTGCTCCAGATCTGCCCACTCC 60

Search completed: January 15, 2005, 20:36:09
 Job time : 172.5 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2005, 17:40:52 ; Search time 722.8 Seconds
(without alignments)
3925.542 Million cell updates/sec

Title: US-10-057-136-6

Perfect score: 60

Sequence: 1 GGGTCAACAGCTCTCCGCGC.....CAGATACCTCGCCAGCTCCA 60

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 2364489745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl: *
1: gb_ba: *
2: gb_hhg: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pac: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_slc: *
12: gb_sy: *
13: gb_un: *
14: gb_vl: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	2297	6	BD272907 A recombi
2	33.6	56.0	176263	2	AC102389 Mus muscu
3	33.6	56.0	223290	2	AC140409 Mus muscu
4	31.6	52.7	525	6	BD225141 Medicinal
5	31.6	52.7	891	6	BD225147 Medicinal
6	31.6	52.7	1371	6	BD225144 Medicinal
7	31.6	52.7	1737	6	BD225150 Medicinal
8	30.8	51.3	149500	2	AC138636 leishman
9	29.6	49.3	93	6	AX703426 Sequence
10	29.6	49.3	157	6	AX703428 Sequence
11	29.6	49.3	82400	10	AC090495 Genomic s
12	29.6	49.3	164629	10	AC121094 Genomic s
13	29.6	49.3	174264	10	AL591373 Mouse DNA
14	29.6	49.3	207182	10	AC144938 Mus muscu
15	29.6	49.3	219963	2	AC124587 Mus muscu
16	29.6	49.3	223724	10	AC114003 Mus muscu
17	29	48.3	1638	10	BC026670 Mus muscu
18	29	48.3	2446	4	OCU37769 Oryctolagus
19	29	48.3	2524	10	BC058977 Mus muscu

c	20	28.8	48.0	249767	10	AC114539	AC114539 Mus muscu
	21	28.6	47.7	519	6	AR316604	AR316604 Sequence
	22	28.6	47.7	534	6	AR316608	AR316608 Sequence
	23	28.6	47.7	40666	2	AC019903	AC019903 Drosophila
	24	28.6	47.7	86391	3	AC004716	AC004716 Drosophila
	25	28.6	47.7	171000	3	AC092244	AC092244 Drosophila
	26	28.6	47.7	192707	2	AC116250	AC116250 Rattus no
	27	28.6	47.7	195798	2	AC118890	AC118890 Rattus no
	28	28.6	47.7	337636	3	AE003584	AE003584 Drosophila
	29	28.4	47.3	110000	8	CR382131	CR382131 Rattus no
	30	28.2	47.0	116409	10	AC131118	AC131118 Mus muscu
	31	28.2	47.0	178544	10	AC124557	AC124557 Mus muscu
	32	28.2	47.0	213037	2	AC093986	AC093986 Rattus no
	33	28.2	47.0	233411	2	AC109951	AC109951 Rattus no
	34	28.2	47.0	234248	2	AC119385	AC119385 Rattus no
	35	28.2	47.0	259308	2	AC111121	AC111121 Mus muscu
	36	28	46.7	110000	2	AC091341	AC091341 Rattus no
	37	28	46.7	124244	10	AL929026	AL929026 Mouse DNA
	38	28	46.7	144093	10	AC109193	AC109193 Mus muscu
	39	28	46.7	157092	10	AL805911	AL805911 Mouse DNA
	40	28	46.7	194209	10	AL671894	AL671894 Mouse DNA
	41	28	46.7	199993	10	AC102783	AC102783 Mus muscu
	42	28	46.7	200574	10	BX005304	BX005304 Mouse DNA
	43	28	46.7	202311	2	AC116769	AC116769 Mus muscu
	44	28	46.7	205054	10	AC098712	AC098712 Mus muscu
	45	28	46.7	225238	2	AC097825	AC097825 Rattus no

ALIGNMENTS

RESULT 1	BD272907	2297 bp	DNA	linear	PAT 17-JUL-2003
LOCUS	BD272907				
DEFINITION	A recombinant vector expressing multiple constitutively molecules and uses thereof.				
ACCESSION	BD272907	1	GI:33082675		
VERSION	BD272907	1	GI:33082675		
KEYWORDS	JP 2002531133-A/1.				
SOURCE	synthetic construct				
ORGANISM	artificial sequences.				
REFERENCE	1 (bases 1 to 2297)				
AUTHORS	Schlom, J., Hodge, J., and Pantcali, D.				
TITLE	A recombinant vector expressing multiple constitutively molecules and uses thereof				
JOURNAL	Patent: JP 2002531133-A 1 24-SEP-2002;				
COMMENT	THE UNITED STATES OF AMERICA, THERION BIOLOGICS CORP				
	OS Artificial Sequence				
	PN JP 2002531133-A/1				
	PD 24-SEP-2002				
	PF 12-NOV-1999 JP 2000586927				
	PI 09-DEC-1998 US 60/111582				
	PR JEFFREY SCHLOM, JAMES HODGE, DENNIS PANICALI				
	PC C12N15/02, A61K35/12, A61K35/74, A61K35/76, A61K38/00, A61K39/00,				
	PC A61K39/12, A61K39/21, A61K39/235, A61K39/245, A61K39/275, A61K39/				
	PC A61K39/125, A61K39/21, A61K39/235, A61K39/245, A61K39/275, A61K39/				
	PC 29, A61K48/00,				
	PC A61P1/04, A61P31/04, A61P31/10, A61P31/12, A61P35/00, A61P37/02, PC				
	A61P37/06,				
	PC C12N1/15, C12N1/19, C12N1/21, C12N6/10, C12N7/00, C12Q1/02, G01N3/				
	PC C12N5/00, A61K37/02				
	CC Description of Artificial Sequence: VECTOR SEQUENCE FH Key				
	FT source				
	1. .2297				
	Location/Qualifiers				
	1. .2297				
	/organism="artificial sequence".				
	/mol_type="synthetic construct"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:32630"				
	ORIGIN				

Query Match 100.0%; Score 60; DB 6; Length 2297;
 Best Local Similarity 100.0%; Pred. No. 1,8e-09;
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGTCACAGCTCTCCCGCATGCGGGGTTACTTCTGCTCAGATATCTCGCCAGCTCCCA 60
 |||||
 Db 526 GGGTCACAGCTCTCTCCCGCATGCGGGGTTACTTCTGCTCAGATATCTCGCCAGCTCCCA 585

RESULT 2
 AC102389/c
 LOCUS AC102389
 DEFINITION Mus musculus clone RP24-374N16, WORKING DRAFT SEQUENCE, 12
 unordereed pieces.
 AC102389
 AC102389.3 GI:28893769
 HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 176263)
 Birren, B., Nusbaum, C. and Lander, E.
 Mus musculus, clone RP24-374N16
 Unpublished
 2 (bases 1 to 176263)
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
 Anderson, S., Barna, N., Bastien, V., Boguslavskiy, L., Bouhgalter, B.,
 Brown, A., Camarata, J., Campoliano, A., Chang, J., Chazaro, B.,
 Choepel, Y., Colangelo, M., Collins, S., Collamore, A., Cook, A.,
 Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S.,
 Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
 Gande, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, P.,
 Hagos, B., Hearford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,
 Jones, C., Kamat, A., Karatas, A., Kells, C., Lacroque, K.,
 Lamas, A., Lander, E., Lander, T., Lehoczy, J., Levine, R., Liu, G.,
 Maclean, C., Macdonald, P., Major, J., Margulis, N., Matthews, C.,
 McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrum, J.,
 Menus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C.,
 Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D.,
 Oliver, J., Peterson, K., Phukhang, P., Pierre, N., Pollara, V.,
 Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
 Roman, J., Roselli, M., Roy, A., Santos, R., Schauer, S., Schupack, R.,
 Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
 Strauss, N., Subramanian, A., Talamas, J., Testfaye, S., Theodore, J.,
 Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
 Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 176263)
 Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,
 Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T.,
 Boguslavskiy, L., Bouhgalter, B., Camarata, J., Chang, J., Choepel, Y.,
 Collamore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K.,
 Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S.,
 Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S.,
 Graham, L., Grand-Pierre, N., Hagez, N., Hagopian, D., Hagos, B.,
 Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
 Kamat, A., Karatas, A., Kells, C., Lander, E., Levine, R.,
 Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., Maclean, C.,
 Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,
 Meldrum, J., Menus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J.,
 Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
 O'Neil, D., Oliver, J., Peterson, K., Phukhang, P., Pierre, N.,
 Rachupka, A., Ramsamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P.,
 Roman, J., Schauer, S., Schupack, R., Seaman, S., Severy, P., Smith, C.,
 Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, J., Talamas, J.,
 Testfaye, S., Theodore, J., Topham, K., Travers, M.,
 Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X.,
 Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
 JOURNAL
 COMMENT

Direct Submission
 Submitted (09-MAR-2003) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Mar 9, 2003 this sequence version replaced gi:22381003.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

 Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu

 Project Information
 Center project name: 374 N 16
 Center clone name: 118673

 Summary Statistics
 Sequencing vector: Plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye, 1000 of reads
 Assembly program: Phred, version 0.960731
 Consensus quality: 173405 bases at least Q40
 Consensus quality: 174612 bases at least Q30
 Consensus quality: 174923 bases at least Q20
 Insert size: 170000; agarose-fp
 Insert size: 175163; sum-of-contigs
 Quality coverage: 9.5 in Q20 bases; agarose-fp
 Quality coverage: 9.2 in Q20 bases; sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 12 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 1859: contig of 1859 bp in length
 * 1860 1959: gap of 100 bp
 * 1960 3317: contig of 1358 bp in length
 * 3318 3417: gap of 100 bp
 * 3418 4588: contig of 1171 bp in length
 * 4589 4688: gap of 100 bp
 * 4689 6418: contig of 1730 bp in length
 * 6419 6518: gap of 100 bp
 * 6519 10148: contig of 3630 bp in length
 * 10149 10248: gap of 100 bp
 * 10249 15378: contig of 5128 bp in length
 * 15379 15476: gap of 100 bp
 * 15477 26015: contig of 10539 bp in length
 * 26016 26115: gap of 100 bp
 * 26116 40612: contig of 14497 bp in length
 * 40613 40712: gap of 100 bp
 * 40713 58063: contig of 17351 bp in length
 * 58064 58163: gap of 100 bp
 * 58164 82401: contig of 24238 bp in length
 * 82402 82501: gap of 100 bp
 * 82502 140587: contig of 58066 bp in length
 * 140588 140687: gap of 100 bp
 * 140688 176263: contig of 35576 bp in length.

FEATURES
 source
 1.176263
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /clone="RP24-374N16"
 /clone_1lb="RP24-374N16"
 1.1859
 /note="assembly_fragment
 clone_end:SP6
 vector_side:left"
 1960..3317
 /note="assembly_fragment"
 3418..4588
 /note="assembly_fragment"

misc_feature
 misc_feature
 misc_feature

```

misc_feature      4689..6418
                  /note="assembly_fragment"
misc_feature      6519..10148
                  /note="assembly_fragment"
misc_feature      10249..15376
                  /note="assembly_fragment"
misc_feature      15477..26015
                  /note="assembly_fragment"
misc_feature      26116..40612
                  /note="assembly_fragment"
misc_feature      40713..58063
                  /note="assembly_fragment"
misc_feature      58164..82401
                  /note="assembly_fragment"
misc_feature      82502..140587
                  /note="assembly_fragment"
misc_feature      140688..176263
                  /note="assembly_fragment"
                  clone_end:T7
                  vector_side:right"

```

ORIGIN

```

Query Match      56.0%; Score 33.6; DB 2; Length 176263;
Best Local Similarity 75.0%; Pred. No. 1.9;
Matches 42; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

```

```

OY      5 CACAGCTCTCCGCTCATGGGCTTCTTCTGCTCAGATCTCCGCCAGCTCCA 60
DB      36031 CAGTGCTCTCCCGAGCTCCAGTGTTCCCGCAGCTCCAGATGCTCCCGAGCTCCA 35976

```

```

RESULT 3
AC140409      223290 bp      DNA      linear      HTG 23-FEB-2003
LOCUS

```

```

AC140409      Mus musculus chromosome UNK clone RP23-138K22, WORKING DRAFT
SEQUENCE, 16 unordered pieces.

```

```

AC140409.1 GT:28475573
AC140409      HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS      Mus musculus (house mouse)
SOURCE

```

```

ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```

```

REFERENCE
AUTHORS      McPherson,J.D. and Waterston,R.H.
TITLE        The sequence of Mus musculus clone
JOURNAL      Unpublished

```

```

AUTHORS      McPherson,J.D. and Waterston,R.H.
REFERENCE    Direct Submision
TITLE        Submitted (23-FEB-2003) Genome Sequencing Center, 4444 Forest Park
JOURNAL      Parkway, St. Louis, MO 63108, USA

```

COMMENT

```

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@watsn.wustl.edu
----- Project Information -----
Center project name: M_BA0138K22
----- Summary Statistics -----
Sequencing vector: M13; 0%
Sequencing vector: plasmid; 100%
Chemistry: Dye-primer ET; 0% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 218110 bases at least Q40
Consensus quality: 218643 bases at least Q30
Consensus quality: 219019 bases at least Q20
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 16 contigs. The true order of the pieces

```

FEATURES

source

```

misc_feature      1..1248
                  /note="assembly_name:Contig15"
misc_feature      1349..4137
                  /note="assembly_name:Contig18"
misc_feature      4238..8834
                  /note="assembly_name:Contig19"
misc_feature      8935..14846
                  /note="assembly_name:Contig20"
misc_feature      14947..23063
                  /note="assembly_name:Contig21"
misc_feature      23164..32288
                  /note="assembly_name:Contig22"
misc_feature      32389..44729
                  /note="assembly_name:Contig23"
misc_feature      44830..61526
                  /note="assembly_name:Contig24"
misc_feature      61627..72386
                  /note="assembly_name:Contig25"
misc_feature      72487..85071
                  /note="assembly_name:Contig26"
misc_feature      85172..98342
                  /note="assembly_name:Contig27"
misc_feature      98443..114309
                  /note="assembly_name:Contig28"
misc_feature      114410..141127
                  /note="assembly_name:Contig29"
misc_feature      141228..165507
                  /note="assembly_name:Contig30"
misc_feature      165608..194683

```

```

* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1      1248: contig of 1248 bp in length
1348: gap of unknown length
4137: contig of 2789 bp in length
4138: gap of unknown length
4238: contig of 4537 bp in length
8835: gap of unknown length
8935: contig of 5912 bp in length
14847: gap of unknown length
14946: contig of 8117 bp in length
23063: contig of 8117 bp in length
23163: gap of unknown length
23164: contig of 9125 bp in length
32288: contig of 9125 bp in length
32389: gap of unknown length
44729: contig of 12341 bp in length
44829: gap of unknown length
44730: gap of unknown length
44830: contig of 16697 bp in length
61526: gap of unknown length
61527: gap of unknown length
72386: contig of 10760 bp in length
72387: gap of unknown length
72486: gap of unknown length
72487: contig of 12585 bp in length
85071: gap of unknown length
85171: gap of unknown length
98342: contig of 13171 bp in length
98343: gap of unknown length
98442: gap of unknown length
114309: contig of 15667 bp in length
114310: gap of unknown length
141127: contig of 26718 bp in length
141227: gap of unknown length
141228: gap of unknown length
165507: contig of 24280 bp in length
165607: gap of unknown length
165608: contig of 29076 bp in length
194684: gap of unknown length
194784: 223290: contig of 28507 bp in length.
Location/Qualifiers
1..223290
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="UNK"
/clone="RP23-138K22"
1..1248
/note="assembly_name:Contig15"
1349..4137
/note="assembly_name:Contig18"
4238..8834
/note="assembly_name:Contig19"
8935..14846
/note="assembly_name:Contig20"
14947..23063
/note="assembly_name:Contig21"
23164..32288
/note="assembly_name:Contig22"
32389..44729
/note="assembly_name:Contig23"
44830..61526
/note="assembly_name:Contig24"
61627..72386
/note="assembly_name:Contig25"
72487..85071
/note="assembly_name:Contig26"
85172..98342
/note="assembly_name:Contig27"
98443..114309
/note="assembly_name:Contig28"
114410..141127
/note="assembly_name:Contig29"
141228..165507
/note="assembly_name:Contig30"
165608..194683

```

misc_feature		/note="assembly_name:Contig31" 194784 /note="assembly_name:Contig32"	
ORIGIN			
Query Match	56.0%; Score 33.6; DB 2; Length 223290;		
Best Local Similarity	75.0%; Pred. No. 1.9;		
Matches	42; Conservative 0; Mismatches 14; Indels 0; Gaps 0;		
QY	5 CAACAGCTCTCCCTCCGCTCATGGGGTTACTTCTGCTCCAGATACTCGCCAGCTCCA 60		
DB	111887 CAGGTGCTCCCTCCGCTCATGGGGTTACTTCCCGCCAGCTCCAGATGCTTCCCGAGCTCCA 111942		
RESULT 4			
BD225141			
LOCUS	BD225141	525 bp	linear PAT 17-JUL-2003
DEFINITION	Medicinal composition having antitumor effect and containing DNA encoding antigenic protein.		
ACCESSION	BD225141		
VERSION	BD225141.1 GI:33034911		
KEYWORDS	JP 2002528519-A/2.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 525)		
TITLE	Pallente, D., Massimo, A.M.D. and Desantis, R.		
JOURNAL	Medicinal composition having antitumor effect and containing DNA encoding antigenic protein		
COMMENT	Patent: JP 2002528519-A 2 03-SEP-2002;		
	MEARINI RICERCH SPA		
	OS Homo sapiens (human)		
	PN JP 2002528519-A/2		
	PD 03-SEP-2002		
	PF 18-OCT-1998 JP 2000579265		
	PR 30-OCT-1998 IT MI98A002330		
	PI DINO PALLENTE, ANNA MARIA D MASSIMO, RITA DESANTIS PC		
	A61K38/00, A61K35/76, A61K39/00, A61K48/00, A61P35/00, C12N15/09, PC		
	A61K37/02,		
	PC C12N15/00		
	CC Medicinal composition having antitumor effect and containing		
	CC DNA encoding		
	CC antigenic protein		
	CH Key		
	FT source		
FEATURES			
source	1..525		
	Location/Qualifiers		
	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9606"		
ORIGIN			
Query Match	52.7%; Score 31.6; DB 6; Length 525;		
Best Local Similarity	74.1%; Pred. No. 5.4;		
Matches	40; Conservative 0; Mismatches 14; Indels 0; Gaps 0;		
QY	7 ACAGCTCTCTCCGCTCATGGGGTTACTTCTGCTCCAGATACTCGCCAGCTCCA 60		
DB	262 ACTGCTCCACGACAGCAGGTGTTACTCTCGGCTCCGGATACAGGCCGGGCCCA 315		
RESULT 5			
BD225147			
LOCUS	BD225147	891 bp	DNA linear PAT 17-JUL-2003
DEFINITION	Medicinal composition having antitumor effect and containing DNA encoding antigenic protein.		
ACCESSION	BD225147		
VERSION	BD225147.1 GI:33034917		
KEYWORDS	JP 2002528519-A/8.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		

REFERENCE	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE	1 (bases 1 to 891)
JOURNAL	Pallente,D., Massimo,A.M.D. and Desantis,R. Medicinal composition having antitumor effect and containing DNA encoding antigenic protein Patent: JP 2002528519-A 8 03-SEP-2002; MENARINI RICERCHE SPA
COMMENT	OS Homo sapiens (human) PN JP 2002528519-A/8 PD 03-SEP-2002 PF 18-OCT-1999 JP 2000579265 PR 30-OCT-1998 IT MI98A002330 PI DINO PALLENTE,ANNA MARIA D MASSIMO,RITA DESANTIS PC A61K38/00,A61K35/76,A61K39/00,A61K48/00,A61P35/00,C12N15/09,PC A61K37/02, PC C12N15/00 CC Medicinal composition having antitumor effect and containing CC DNA encoding CC antigenic protein FH Key Location/Qualifiers FT source 1..891 /organism='Homo sapiens (human)'. FT Location/Qualifiers 1..891 /organism='Homo sapiens (human)'.
FEATURES	source
ORIGIN	Location/Qualifiers 1..891 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606"
Query Match	52.7%; Score 31.6; DB 6; Length 891;
Best Local Similarity	74.1%; Pred.No.5.6;
Matches	40; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
Dn	628 ACTGCTCCACAGCAGCACGGTGTACTCGGCTCCGGATACCAGCGGCCCA 681
Oy	7 ACAGTCCTCCCGGCTCATNGGGTTACTTGCTGCCTCAGATACTCGCCACGTCCA 60
RESULT 6	BD225144 1371 bp DNA linear PAT 17-JUL-2003
LOCUS	BD225144
DEFINITION	Medicinal composition having antitumor effect and containing DNA encoding antigenic protein.
ACCESSION	BD225144
VERSION	BD225144.1 GI:33034914
KEYWORDS	JP 2002528519-A/5.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 1371)
AUTHORS	Pallente,D., Massimo,A.M.D. and Desantis,R.
TITLE	Medicinal composition having antitumor effect and containing DNA encoding antigenic protein Patent: JP 2002528519-A 5 03-SEP-2002; MENARINI RICERCHE SPA
JOURNAL	OS Homo sapiens (human) PN JP 2002528519-A/5 PD 03-SEP-2002 PF 18-OCT-1999 JP 2000579265 PR 30-OCT-1998 IT MI98A002330 PI DINO PALLENTE,ANNA MARIA D MASSIMO,RITA DESANTIS PC A61K38/00,A61K35/76,A61K39/00,A61K48/00,A61P35/00,C12N15/09,PC A61K37/02, PC C12N15/00 CC Medicinal composition having antitumor effect and containing CC DNA encoding CC antigenic protein FH Key Location/Qualifiers FT source 1..1371 /organism='Homo sapiens (human)'.
COMMENT	location/Qualifiers
FEATURES	location/Qualifiers

source 1.1371
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 52.7%; Score 31.6; DB 6; Length 1371;
Best Local Similarity 74.1%; Pred. No. 5.8;
Matches 40; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 7 ACAGCTCTCCCGCTCATGGGGTTACTTGTGCTCCAGATACCTGCCAGCTCCA 60
|||
331 ACCGCTCCACGACACACGCGTTTACTCGGCTCCGAGTACAGCCGCCCA 384

RESULT 7
BD225150 1737 bp DNA linear PAT 17-JUL-2003
DEFINITION Medicinal composition having antitumor effect and containing DNA
encoding antigenic protein.
ACCESSION BD225150.1 GI:33034920
KEYWORDS JP 2002528519-A/11.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 1737)
Pallente,D., Masstmo,A.M.D. and Desantis,R.
TITLE Medicinal composition having antitumor effect and containing DNA
encoding antigenic protein
Patent: JP 2002528519-A 11 03-SEP-2002;
JOURNAL MENARINI RICERCH SPA
COMMENT OS Homo sapiens (human)
PN JP 2002528519-A/11
PD 03-SEP-2002
PF 18-OCT-1999 JP 2000579265
PR 30-OCT-1998 IT MI98A002330
PI DINO PALLENTE,ANNA MARIA D MASSIMO,RITA DESANTIS PC
A61K38/00,A61K35/76,A61K39/00,A61K48/00,A61P35/00,C12N15/09, PC
A61K37/02,
PC C12N15/00
CC Medicinal composition having antitumor effect and containing
CC CC DNA encoding
CC CC antigenic protein
CC CC Location/Qualifiers
FT Key 1.1737
FT source /organism="Homo sapiens (human)".
FEATURES
source Location/Qualifiers
1.1737
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 52.7%; Score 31.6; DB 6; Length 1737;
Best Local Similarity 74.1%; Pred. No. 6;
Matches 40; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 7 ACAGCTCTCCCGCTCATGGGGTTACTTGTGCTCCAGATACCTGCCAGCTCCA 60
|||
697 ACCGCTCCACGACACACGCGTTTACTCGGCTCCGAGTACAGCCGCCCA 750

RESULT 8
AC138636 149500 bp DNA linear HTG 25-FEB-2003
LOCUS Leishmania major chromosome 29 clone LB00573 strain Friedlin,***
DEFINITION SEQUENCING IN PROGRESS ***, 6 ordered pieces.
ACCESSION AC138636
AC138636.2 GI:28557887
VERSION HTG: HTGS PHASE2.
KEYWORDS Leishmania major
SOURCE

ORGANISM Leishmania major
Eukaryota; Euzlenozoa; Kinetoplastida; Trypanosomatidae;
Leishmania
Leishmania

REFERENCE 1 (bases 1 to 149500)
Myler,P.J., Sisk,E., Vogt,C., Robertson,L., Munden,H., Worthey,L.,
Nelson,S., Ivens,A., Seyler,A., Rinta,J. and Stuart,K.
TITLE Direct Submission
JOURNAL Submitted (14-JAN-2003) Seattle Biomedical Research Institution, 4
Nickerson Street, Seattle, WA 98109-1651, USA
AUTHORS 2 (bases 1 to 149500)
Myler,P.J., Sisk,E., Vogt,C., Robertson,L., Munden,H., Worthey,L.,
Nelson,S., Ivens,A., Seyler,A., Rinta,J. and Stuart,K.
DIRECT SUBMISSION
JOURNAL Submitted (25-FEB-2003) Seattle Biomedical Research Institution, 4
Nickerson Street, Seattle, WA 98109-1651, USA
COMMENT On Feb 25, 2003 this sequence version replaced gi:27733947.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 6 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 17713: contig of 17713 bp in length
* 17714 18016: gap of unknown length
* 18017 21356: contig of 3340 bp in length
* 21357 21658: gap of unknown length
* 21659 26396: contig of 4738 bp in length
* 26397 26698: gap of unknown length
* 26699 72247: contig of 45549 bp in length
* 72248 72549: gap of unknown length
* 72550 123849: contig of 51300 bp in length
* 123850 124151: gap of unknown length
* 124152 149500: contig of 25349 bp in length.
FEATURES
source Location/Qualifiers
1.149500
/organism="Leishmania major"
/mol_type="genomic DNA"
/strain="Friedlin"
/db_xref="taxon:5664"
/chromosome="29"
/clone="LB00573"

ORIGIN

Query Match 51.3%; Score 30.8; DB 2; Length 149500;
Best Local Similarity 70.7%; Pred. No. 16;
Matches 41; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 2 GGTCACAGCTCTCCCGCTCATGGGGTTACTTGTGCTCCAGATACCTGCCAGCTCC 59
|||
DB 116463 GGTGAACAGCTGTACTGCTCACGAGGTGCTGCGCGCTCTCTCCACAACTGC 116520

RESULT 9
AX703426/c 93 bp DNA linear PAT 03-APR-2003
LOCUS AX703426
DEFINITION Sequence 4 from Patent WO02086505.
ACCESSION AX703426
AX703426.1 GI:29538417
VERSION
KEYWORDS AX703426.1 GI:29538417
SOURCE
ORGANISM
synthetic construct
synthetic construct
artificial sequences.
REFERENCE 1
Benson,R.S.
AUTHORS
TITLE Intracellular analysis
JOURNAL Patent: WO 02086505-A 4 31-OCT-2002;
THE VICTORIA UNIVERSITY OF MANCHESTER (GB)
FEATURES
source Location/Qualifiers
1.93
/organism="synthetic construct"

Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., MacDonald, P., Major, J., Margus, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Plerre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schnupack, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strause, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, D., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL
REFERENCE
AUTHORS

Submitted (15-MAY-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 164629)
Anderson, M., Anderson, S., Arachchi, H.M., Barina, N., Baetien, V., Bloom, T., Boguslavsky, L., Boukhalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., Deatellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Fero, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hago, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R., Maclean, C., MacDonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Plerre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schnupack, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL
REFERENCE
AUTHORS

Submitted (24-APR-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 164629)
Birren, B., Nisbaum, C., Landers, E., Aboueleil, A., Allen, N., Anderson, M., Anderson, S., Arachchi, H.M., Barina, N., Baetien, V., Bloom, T., Boguslavsky, L., Boukhalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., Deatellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Fero, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hago, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R., Maclean, C., MacDonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Plerre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schnupack, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL
COMMENT

Submitted (30-JUL-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 30, 2004 this sequence version replaced gi:6518666.
All repeats were identified using RepeatMasker:
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
Genome Center
Center: Whitehead Institute/MIT Center for Genome Research
Center code: WIBR

FEATURES

source

Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@road.mit.edu
----- Project Information
Center project name: L21248
Center clone name: 322_1_3

Some of the sequence contained within base pairs 119380 to the end of the clone was stolen from accession AC144938.
Location/Qualifiers
1..164629
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="18"
/map="18"
/clone="RP24-32213"
/clone_1ib="RPCI-24 Male Mouse BAC"

misc_feature

1..4
/note="clone boundary
clone end:SP6
site:Wb01"
complement(3..179)
/rpt_family="B4A"
complement(1847..2028)
/rpt_family="MIR"
2200..2221
/rpt_family="AT_rich"
/rpt_family="TAA"
4912..4939
/rpt_family="TAA"
5294..5319
/rpt_family="GA"
5797..5844
/rpt_family="TTTTTG"
5961..5995
/rpt_family="GC_rich"
7436..7559
/rpt_family="C-rich"
7630..7660
/rpt_family="A"
7673..7824
/rpt_family="GA"
7871..8023
/rpt_family="TGA"
8027..8070
/rpt_family="CGAG"
10319..10353
/rpt_family="AT_rich"
11283..11312
/rpt_family="TAA"
14721..14888
/rpt_family="CT-rich"
complement(14857..15037)
/note="single clone coverage"
complement(14857..14864)
/note="30 qual single clone coverage"
complement(14884..14890)
/note="30 qual single clone coverage"
14915..15039
/rpt_family="CT-rich"
16447..16631
/rpt_family="TGA"
16480..16647
/note="single clone coverage"
16648..16785
/rpt_family="TGA"
17445..17501
/rpt_family="GA"
17973..18039
/rpt_family="TCA"
18040..18068
/rpt_family="CA"
complement(19489..19649)
repeat_region

```

repeat_region      /rpt_family="MLT1J2"
                    25239..25264
                    /rpt_family="(CA)n"
repeat_region      25346..25389
                    /rpt_family="(CA)n"
repeat_region      25390..25416
                    /rpt_family="(CA)n"
repeat_region      26450..26510
                    /rpt_family="(TG)n"
repeat_region      26926..26995
                    /rpt_family="(CAGAGA)n"
repeat_region      27650..27790
                    /rpt_family="L2"
repeat_region      27921..27954
                    /rpt_family="(CA)n"
repeat_region      28145..28311
                    /rpt_family="MIR"
repeat_region      29591..29696
                    /rpt_family="MIR"
repeat_region      complement(29731..29889)
                    /rpt_family="MIR"
repeat_region      complement(30334..30776)
                    /rpt_family="L2"
repeat_region      complement(32737..32962)
                    /rpt_family="B4A"
repeat_region      33251..33481
                    /rpt_family="MIR"
repeat_region      33884..34054
                    /rpt_family="RSINE1"
repeat_region      complement(34313..34470)
                    /rpt_family="B1F"

```

Query Match 49.3%; Score 29.6; DB 10; Length 164629;
 Best Local Similarity 73.1%; Pred. No. 42;
 Matches 38; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 8 CAGCTCCCTCCGCTCATGGGTTACTTCTGCTCCAGATCTCGCCAGCTCC 59
 DB 153766 CAGCTCCCTCCGCTCATGGGTTACTTCTGCTCCAGATCTCGCCAGCTCC 153715

RESULT 13
 AL591373 174264 bp DNA linear ROD 29-JUN-2002
 LOCUS Mouse sequence from clone RP23-268M6 on chromosome 18, complete
 DEFINITION
 ACCESSION AL591373 GI:15983933
 VERSION AL591373.8
 KEYWORDS HTG.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 174264)
 Philimore, B.
 Direct Submission
 Submitted (29-JUN-2002) Wellcome Trust Sanger Institute, Hinxton,
 humbrey@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
 On Oct 8, 2001 this sequence data is compared from overlapping clones.
 During sequence assembly finished as follows unless otherwise noted: all
 regions were either double-stranded or sequenced with an alternate
 chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one plasmid subclone or more than one M13 subclone; and the
 assembly was confirmed by restriction digest. The following

FEATURES
 source
 Location/Qualifiers
 1..174264
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="18"
 /clone_id="RP23-268M6"
 /clone_1id="RPC1-23"

ORIGIN
 Query Match 49.3%; Score 29.6; DB 10; Length 174264;
 Best Local Similarity 73.1%; Pred. No. 42;
 Matches 38; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 8 CAGCTCCCTCCGCTCATGGGTTACTTCTGCTCCAGATCTCGCCAGCTCC 59
 DB 111108 CAGCTCCCTCCGCTCATGGGTTACTTCTGCTCCAGATCTCGCCAGCTCC 111159

RESULT 14
 AC144938 207182 bp DNA linear ROD 15-MAY-2004
 LOCUS Mus musculus BAC clone RP24-129H18 from chromosome 18, complete
 DEFINITION
 ACCESSION AC144938 GI:45237301
 VERSION AC144938
 KEYWORDS HTG.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 207182)
 Sweatengen-Shahid, S., Shahid, S., Bielicki, L. and Hakenson, W.
 The sequence of Mus musculus BAC clone RP24-129H18
 2 (bases 1 to 207182)
 Wilson, R.K.
 Direct Submission
 Submitted (27-MAY-2003) Genome Sequencing Center, 4444 Forest Park
 Parkway, St. Louis, MO 63108, USA
 3 (bases 1 to 207182)
 Wilson, R.K.
 Direct Submission
 Submitted (25-JUN-2003) Genome Sequencing Center, 4444 Forest Park
 Parkway, St. Louis, MO 63108, USA
 4 (bases 1 to 207182)
 Wilson, R.K.
 Direct Submission
 Submitted (06-MAR-2004) Genome Sequencing Center, 4444 Forest Park
 Parkway, St. Louis, MO 63108, USA
 5 (bases 1 to 207182)
 Wilson, R.K.
 Direct Submission
 Submitted (15-MAY-2004) Genome Sequencing Center, 4444 Forest Park
 Parkway, St. Louis, MO 63108, USA
 On Mar 6, 2004 this sequence version replaced gi:32189675.
 ----- Genome Center
 Center: Washington University Genome Sequencing Center

Center code: WUGSC
 Web site: <http://genome.wustl.edu>
 Contact: submissions@watson.wustl.edu
 Summary Statistics
 Center project name: M_BB0129H18

NOTICE:

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone, fosmid clone or direct clone walk sequence. Sequence from the Mouse Genome Sequencing Consortium whole genome shotgun may have been used to obtain the consensus sequence; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wea Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION:

The RPI-24 BAC library has been constructed by Pieter de Jong and coworkers (<http://www.chori.org>) from male C57BL/6J mouse spleen and/or brain genomic DNA. The clone and detailed information can be obtained from Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone.

FEATURES
 source
 1..207182

/organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="18"
 /map="18"

clone="RP24-129H18"
 /clone_1ib="RPI-24"
 17051..17219

unSURE
 /note="Unresolved simple sequence repeat."

misc_feature
 51528..52399
 /note="CpG_Island (%GC=71.6, o/e=0.81, #CpGs=79)"

misc_feature
 105737..105746
 /note="Sequence derived from PCR product of project DNA."

misc_feature
 128471..128900
 /note="Sequence derived from PCR product of project DNA."
 146793..147510
 /note="Unresolved simple sequence repeat."

ORIGIN

Query Match 49.3% Score 29.6; DB 10; Length 207182;
 Best Local Similarity 73.1% Pred. No. 43;
 Matches 38; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 8 CAGCTCTCCGCGCTCATGGGTTACTTCTGCTCCAGATCTGCCAGCTCC 59
 |||||
 Db 172796 CAGCTCTCCGCGCTCATGGGTTACTTCTGCTCCAGATCTGCCAGCTCC 59
 |||||

RESULT 15
 AC124587 219963 bp DNA linear HTG 10-JUL-2004
 LOCUS AC124587
 DEFINITION Mus musculus chromosome 1 clone RP23-132G24, WORKING DRAFT
 SEQUENCE 5 unordered pieces.
 AC124587
 AC124587.4 GI:50201932
 VERSION AC124587.4
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 219963)
 Wilson,R.K.
 The sequence of Mus musculus clone
 Unpublished
 2 (bases 1 to 219963)
 McPherson,J.D. and Waterston,R.H.
 Direct Submission
 Submitted (14-JUN-2002) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
 3 (bases 1 to 219963)
 Wilson,R.K.
 Direct Submission
 Submitted (10-JUL-2004) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
 On Jul 10, 2004 this sequence version replaced gi:22475935.

Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site:<http://genome.wustl.edu>
 Contact: submissions@watson.wustl.edu
 Project Information
 Center project name: M_BA0132G24

Summary Statistics

Sequencing vector: M13; 0%
 Sequencing vector: Plasmid; 100%
 Chemistry: Dye-primer ET; 0% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 217311 bases at least Q40
 Consensus quality: 217860 bases at least Q30
 Consensus quality: 218107 bases at least Q20

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 5 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 287: contig of 287 bp in length
 288 387: gap of unknown length
 388 9268: contig of 8881 bp in length
 9269 9369: gap of unknown length
 9369 36735: contig of 27367 bp in length
 36736 36835: gap of unknown length
 36836 76965: contig of 40130 bp in length
 76966 77065: gap of unknown length
 77066 219963: contig of 142898 bp in length.

FEATURES

source

1..219963
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="1"
 /clone="RP23-132G24"

misc_feature

1..287
 /note="assembly_name:Contig36
 clone_end:T7
 vector_side:right"

misc_feature

388..9268
 /note="assembly_name:Contig38"

misc_feature

9369..36735
 /note="assembly_name:Contig39"

misc_feature

36836..76965
 /note="assembly_name:Contig40"
 77066..219963
 /note="assembly_name:Contig41
 clone_end:SP6

Db 1 GGATCCACCGCGCGCTGCGCACGAGTACGTCCGCCGCCGACACGCGCCCGCTGCC 60

RESULT 2

US-10-406-317-41
; Sequence 41, Application US/10406317
; Publication No. US20040019195A1
; GENERAL INFORMATION:
; APPLICANT: Schlom, Jeffrey;
; APPLICANT: Hodge, James;
; APPLICANT: Panicali, Dennis;
; TITLE OF INVENTION: A recombinant vector expressing multiple constitutively
; FILE REFERENCE: 38163-0189
; CURRENT APPLICATION NUMBER: US/10/406,317
; CURRENT FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: US/09/856,988
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: PCT/US99/26866
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/111,582
; PRIOR FILING DATE: 1998-12-09
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41
; LENGTH: 2297
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: VECTOR
US-10-406-317-41

Query Match 100.0%; Score 60; DB 16; Length 2297;
Best Local Similarity 100.0%; Pred. No. 1.9e-11;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGATCCACCGCGCGCTGCGCACGAGTACGTCCGCCGCCGACACGCGCCCGCTGCC 60
Db 466 GGATCCACCGCGCGCTGCGCACGAGTACGTCCGCCGCCGACACGCGCCCGCTGCC 525

RESULT 3

US-10-057-136-11
; Sequence 11, Application US/10057136
; Publication No. US20030021770A1
; GENERAL INFORMATION:
; APPLICANT: SCHLOM, JEFFREY
; APPLICANT: KANTOR, JUDITH
; APPLICANT: KUFE, DONALD
; APPLICANT: PANICALI, DENNIS
; APPLICANT: GRITZ, LINDA
; TITLE OF INVENTION: RECOMBINANT FOX VIRUS FOR IMMUNIZATION AGAINST MUC1
; FILE REFERENCE: 700953/47113C
; CURRENT APPLICATION NUMBER: US/10/057,136
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/366,670
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: PCT/US98/03693
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/038,253
; PRIOR FILING DATE: 1997-02-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-136-11

Query Match 63.7%; Score 38.2; DB 14; Length 60;
Best Local Similarity 78.0%; Pred. No. 0.001;

Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 1 GGATCCACCGCGCGCTGCGCACGAGTACGTCCGCCGCCGACACGCGCCCGCTGCC 59
Db 1 GGTTCACCGCACCTCCGACGAGTACGTCTGCACCCGACACCCGCTCCAGCTCC 59

RESULT 4

US-10-057-136-2
; Sequence 2, Application US/10057136
; Publication No. US20030021770A1
; GENERAL INFORMATION:
; APPLICANT: SCHLOM, JEFFREY
; APPLICANT: KANTOR, JUDITH
; APPLICANT: KUFE, DONALD
; APPLICANT: PANICALI, DENNIS
; APPLICANT: GRITZ, LINDA
; TITLE OF INVENTION: RECOMBINANT FOX VIRUS FOR IMMUNIZATION AGAINST MUC1
; FILE REFERENCE: 700953/47113C
; CURRENT APPLICATION NUMBER: US/10/057,136
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/366,670
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: PCT/US98/03693
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/038,253
; PRIOR FILING DATE: 1997-02-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-136-2

Query Match 58.3%; Score 35; DB 14; Length 60;
Best Local Similarity 74.6%; Pred. No. 0.013;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 1 GGATCCACCGCGCGCTGCGCACGAGTACGTCCGCCGCCGACACGCGCCCGCTGCC 59
Db 1 GGTCACCGCGCGCGCTGCGCACGAGTACGTCTGCACCCGACACCCGCTCCAGCTCC 59

RESULT 5

US-10-447-839A-75/c
; Sequence 75, Application US/10447839A
; Publication No. US20040018181A1
; GENERAL INFORMATION:
; APPLICANT: Kufe, Donald W.
; APPLICANT: Khairpanda, Surender
; APPLICANT: Weithan, Steven D.
; TITLE OF INVENTION: MUC1 INTERFERENCE RNA COMPOSITIONS AND METHODS DERIVED THEREFROM
; FILE REFERENCE: 1000.05.009
; CURRENT APPLICATION NUMBER: US/10/447,839A
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: 10/293,391
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 09/951,938
; PRIOR FILING DATE: 2001-09-11
; PRIOR APPLICATION NUMBER: 60/231,841
; PRIOR FILING DATE: 2000-09-11
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: PatentIn Version 3.2
; SEQ ID NO 75
; LENGTH: 1424
; TYPE: RNA
; ORGANISM: ARTIFICIAL
; FEATURE:
; OTHER INFORMATION: Synthesized Sequence
US-10-447-839A-75

Query Match 63.7%; Score 38.2; DB 14; Length 60;
Best Local Similarity 78.0%; Pred. No. 0.001;

Query Match 58.3%; Score 35; DB 16; Length 1424;
Best Local Similarity 74.6%; Pred. No. 0.0061;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 1 GGATCACCAGCGCGCTGCGACGAGTGCCTGCGCGCCGACACGCGCCCGCTCC 59
Db 1041 GGCTCACCAGCGCGCCCGCCAGCCAGGTGTCACTCGCGCCCGGACACAGCGCGCCCG 983

RESULT 6

US-10-447-839A-20
; Sequence 20, Application US/10447839A
; Publication No. US20040018181A1
; GENERAL INFORMATION:
; APPLICANT: Kufe, Donald W.
; APPLICANT: Khatibanda, Surender
; APPLICANT: Weltman, Steven D.
; TITLE OF INVENTION: MUC1 INTERREFERENCE RNA COMPOSITIONS AND METHODS DERIVED THEREFROM
; FILE REFERENCE: 1000.05.009
; CURRENT APPLICATION NUMBER: US/10/447,839A
; PRIOR FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: 10/293,391
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 09/951,938
; PRIOR FILING DATE: 2001-09-11
; PRIOR APPLICATION NUMBER: 60/231,841
; PRIOR FILING DATE: 2000-09-11
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 20
; LENGTH: 1428
; TYPE: RNA
; ORGANISM: RNA
US-10-447-839A-20

Query Match 58.3%; Score 35; DB 16; Length 1428;
Best Local Similarity 69.5%; Pred. No. 0.0061;
Matches 41; Conservative 3; Mismatches 15; Indels 0; Gaps 0;

Qy 1 GGATCACCAGCGCGCTGCGACGAGTGCCTGCGCGCCGACACGCGCCCGCTCC 59
Db 385 GGCTCACCAGCGCGCCCGCCAGCCAGGTGTCACTCGCGCCCGGACACAGCGCGCCCG 443

RESULT 7

US-10-057-136-19
; Sequence 19, Application US/10057136
; Publication No. US20030021770A1
; GENERAL INFORMATION:
; APPLICANT: SCHLOM, JEFFREY
; APPLICANT: KANTOR, JUDITH
; APPLICANT: Kufe, Donald
; APPLICANT: PANICALI, DENNIS
; APPLICANT: GRITZ, LINDA
; TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
; TITLE OF INVENTION: TUMOR-ASSOCIATED ANTIGEN
; FILE REFERENCE: 700953/47113C
; CURRENT APPLICATION NUMBER: US/10/057,136
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/366,670
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: PCT/US98/03693
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/038,253
; PRIOR FILING DATE: 1997-02-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 1527
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-136-19

Query Match 58.3%; Score 35; DB 14; Length 1527;
Best Local Similarity 74.6%; Pred. No. 0.0061;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 1 GGATCACCAGCGCGCTGCGACGAGTGCCTGCGCGCCGACACGCGCCCGCTCC 59
Db 226 GGCTCACCAGCGCGCCCGCCAGCCAGGTGTCACTCGCGCCCGGACACAGCGCGCCCG 284

RESULT 8

US-09-864-864-280
; Sequence 280, Application US/09864864
; Patent No. US20020102679A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Secrist, Heather
; APPLICANT: Lodes, Michael J.
; APPLICANT: Algate, Paul A.
; APPLICANT: Fling, Steve P.
; APPLICANT: Mannion, Jane
; APPLICANT: Benson, Darin R.
; APPLICANT: Carter, Darrick
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.523
; CURRENT APPLICATION NUMBER: US/09/864,864
; CURRENT FILING DATE: 2001-05-23
; NUMBER OF SEQ ID NOS: 341
; SOFTWARE: Corixa Invention Disclosure Database
; SEQ ID NO 280
; LENGTH: 1721
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-864-864-280

Query Match 58.3%; Score 35; DB 9; Length 1721;
Best Local Similarity 74.6%; Pred. No. 0.0059;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 1 GGATCACCAGCGCGCTGCGACGAGTGCCTGCGCGCCGACACGCGCCCGCTCC 59
Db 562 GGCTCACCAGCGCGCCCGCCAGCCAGGTGTCACTCGCGCCCGGACACAGCGCGCCCG 620

RESULT 9

US-09-967-768A-224
; Sequence 224, Application US/09967768A
; Patent No. US20020150877A1
; GENERAL INFORMATION:
; APPLICANT: Augustus, Meena
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signature
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-72
; CURRENT APPLICATION NUMBER: US/09/967,768A
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US/60/236,109
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,034
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,111
; PRIOR FILING DATE: 2000-09-28
; NUMBER OF SEQ ID NOS: 325
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 224
; LENGTH: 1721
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-967-768A-224

Query Match 58.3%; Score 35; DB 9; Length 1721;

Best Local Similarity 74.6%; Pred. No. 0.0059;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 1 GGATCCACCGCGCGCGCTGCGCAGAGTGACCTCGCGCCGACACAGCGCCCGCTCC 59
Db 562 GGCTCCACCGCGCGCGCGAGCCAGCGGTGTACCTCGGCCCGGACACAGCGCGGCC 620

RESULT 10

US-10-247-703-21
Sequence 21, Application US/10247703
Publication No. US2003063597A1
GENERAL INFORMATION:
APPLICANT: Branigan, Patrick
APPLICANT: Goletz, Theresa J
APPLICANT: Knight, David M
APPLICANT: McCarthy, Stephen G
APPLICANT: Scallion, Bernard J
APPLICANT: Snyder, Linda A
TITLE OF INVENTION: CYTOKINE ADJUVANT VACCINES USING TUMOR ANTIGEN ENCODING NUCLEIC ACIDS
TITLE OF INVENTION: CYTOKINE ADJUVANT ENCODING NUCLEIC ACID
FILE REFERENCE: CEN310
CURRENT APPLICATION NUMBER: US/10/247,703
CURRENT FILING DATE: 2002-09-20
PRIOR APPLICATION NUMBER: 60/328,371
PRIOR FILING DATE: 2001-10-10
NUMBER OF SEQ ID NOS: 77
SOFTWARE: Patentin version 3.1
SEQ ID NO 21
LENGTH: 1721
TYPE: DNA
ORGANISM: Homo sapiens
US-10-247-703-21

Query Match 58.3%; Score 35; DB 14; Length 1721;
Best Local Similarity 74.6%; Pred. No. 0.0059;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 1 GGATCCACCGCGCGCGCTGCGCAGAGTGACCTCGCGCCGACACAGCGCCCGCTCC 59
Db 562 GGCTCCACCGCGCGCGCGAGCCAGCGGTGTACCTCGGCCCGGACACAGCGCGGCC 620

RESULT 11

US-10-097-340-211
Sequence 211, Application US/10097340
Publication No. US20030087250A1
GENERAL INFORMATION:
APPLICANT: John MONAHAN
APPLICANT: Manjula GANNAVARAPU
APPLICANT: Sebastian HOERSCH
APPLICANT: Shubhangi KAMATKAR
APPLICANT: Steve G. KOVATS
APPLICANT: Rachel E. MEYERS
APPLICANT: Michael MORRISSEY
APPLICANT: Peter OLANDT
APPLICANT: Ami SEN
APPLICANT: Peter VERIBY
APPLICANT: Gordon B. MILLS
APPLICANT: Robert C. BAST, Jr.
APPLICANT: Karen LU
APPLICANT: Rosemarie SCHWANDT
APPLICANT: Xumei ZHAO
APPLICANT: Karen GRATT
TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
TITLE OF INVENTION: Assessment, Prevention, and Therapy of Ovarian Cancer
FILE REFERENCE: MRI-030
CURRENT APPLICATION NUMBER: US/10/097,340
CURRENT FILING DATE: 2002-03-14
PRIOR APPLICATION NUMBER: 60/276,025
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 60/325,149
PRIOR FILING DATE: 2001-09-26

PRIOR APPLICATION NUMBER: 60/276,026
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 60/324,967
PRIOR FILING DATE: 2001/09/26
PRIOR APPLICATION NUMBER: 60/311,732
PRIOR FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: 60/325,102
PRIOR FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 60/323,580
PRIOR FILING DATE: 2001-09-19
NUMBER OF SEQ ID NOS: 363
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 211
LENGTH: 1721
TYPE: DNA
ORGANISM: Homo sapiens
US-10-097-340-211

Query Match 58.3%; Score 35; DB 14; Length 1721;
Best Local Similarity 74.6%; Pred. No. 0.0059;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 1 GGATCCACCGCGCGCGCTGCGCAGAGTGACGTGCGCGCCGACACAGCGCCCGCTCC 59
Db 562 GGCTCCACCGCGCGCGCGAGCCAGCGGTGTACCTCGGCCCGGACACAGCGCGGCC 620

RESULT 12

US-10-171-311-155
Sequence 155, Application US/10171311
Publication No. US20030087270A1
GENERAL INFORMATION:
APPLICANT: Schlegel, Robert
APPLICANT: Chen, Yan
APPLICANT: Zhao, Xumei
APPLICANT: Monahan, John
APPLICANT: Kamatkar, Shubhangi
APPLICANT: Glatz, Karen
APPLICANT: Gannavarapu, Manjula
APPLICANT: Hoersch, Sebastian
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
FILE REFERENCE: MRI-035
CURRENT APPLICATION NUMBER: US/10/171,311
CURRENT FILING DATE: 2002-06-12
PRIOR APPLICATION NUMBER: US 60/298,159
PRIOR FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/298,155
PRIOR FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/335,936
PRIOR FILING DATE: 2001-11-14
NUMBER OF SEQ ID NOS: 238
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 155
LENGTH: 1721
TYPE: DNA
ORGANISM: Homo sapiens
US-10-171-311-155

Query Match 58.3%; Score 35; DB 14; Length 1721;
Best Local Similarity 74.6%; Pred. No. 0.0059;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 1 GGATCCACCGCGCGCGCTGCGCAGAGTGACGTGCGCGCCGACACAGCGCCCGCTCC 59
Db 562 GGCTCCACCGCGCGCGCGAGCCAGCGGTGTACCTCGGCCCGGACACAGCGCGGCC 620

RESULT 13

US-10-007-926A-58
Sequence 58, Application US/10007926A
Publication No. US20030143539A1

```
/ GENERAL INFORMATION:
/ APPLICANT: BERTUCCI, FRANCOIS
/ APPLICANT: HOUIGATTE, REMI
/ APPLICANT: BIRNBAUM, DANIEL
/ APPLICANT: NGUYEN, CATHERINE
/ APPLICANT: VIENS, PATRICE
/ APPLICANT: FERT, VINCENT
/ TITLE OF INVENTION: GENE EXPRESSION PROFILING OF PRIMARY BREAST CARCINOMAS
/ TITLE OF INVENTION: USING ARRAYS OF CANDIDATE GENES
/ FILE REFERENCE: 1546-R-00
/ CURRENT APPLICATION NUMBER: US/10/007,926A
/ CURRENT FILING DATE: 2001-12-07
/ PRIOR APPLICATION NUMBER: 60/254,090
/ PRIOR FILING DATE: 2000-12-08
/ NUMBER OF SEQ ID NOS: 468
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 58
/ LENGTH: 1721
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: mucin 1, transmembrane (MUC1) gene.
US-10-007-926A-58

Query Match          58.3%; Score 35; DB 15; Length 1721;
Best Local Similarity 74.6%; Pred. No. 0.0059;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 GGATCCACCGCGCGCTGCGCAGGAGTGACGTGCGCGCCGACACGCGCCCGCTCC 59
DB 562 GGCTCCACCGCGCGCCGCGACGCCAGGTGTCACTCGCGCCGACACCGCGCGCC 620

RESULT 14
US-10-029-517-3
/ Sequence 3, Application US/10029517
/ Publication No. US20030148969A1
/ GENERAL INFORMATION:
/ APPLICANT: Kenneth W. Dobie
/ APPLICANT: Susan J. Myers
/ TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
/ FILE REFERENCE: RTS-0352
/ CURRENT APPLICATION NUMBER: US/10/029,517
/ CURRENT FILING DATE: 2001-12-20
/ NUMBER OF SEQ ID NOS: 107
/ SEQ ID NO 3
/ LENGTH: 1721
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (58)...(1605)
US-10-029-517-3

Query Match          58.3%; Score 35; DB 15; Length 1721;
Best Local Similarity 74.6%; Pred. No. 0.0059;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 GGATCCACCGCGCGCTGCGCAGGAGTGACGTGCGCGCCGACACGCGCCCGCTCC 59
DB 562 GGCTCCACCGCGCGCCGCGACGCCAGGTGTCACTCGCGCCGACACCGCGCGCC 620

RESULT 15
US-10-172-118-775
/ Sequence 775, Application US/10172118
/ Publication No. US2003024374A1
/ GENERAL INFORMATION:
/ APPLICANT: Dai, Hongyue
/ APPLICANT: He, Yudong
/ APPLICANT: Lineley, Peter
/ APPLICANT: Mao, Mao
/ APPLICANT: Roberts, Chris
```

```
/ APPLICANT: Van 't Veer, Laura
/ APPLICANT: Van de Vijver, Marc
/ APPLICANT: Bernards, Rene
/ TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
/ FILE REFERENCE: 9301-175-999
/ CURRENT APPLICATION NUMBER: US/10/172,118
/ CURRENT FILING DATE: 2002-06-14
/ PRIOR APPLICATION NUMBER: 60/380,770
/ PRIOR FILING DATE: 2002-05-14
/ NUMBER OF SEQ ID NOS: 2699
/ SEQ ID NO 775
/ LENGTH: 1721
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ PUBLICATION INFORMATION:
/ DATABASE ACCESSION NUMBER: NM_002456
/ DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-775

Query Match          58.3%; Score 35; DB 15; Length 1721;
Best Local Similarity 74.6%; Pred. No. 0.0059;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 GGATCCACCGCGCGCTGCGCAGGAGTGACGTGCGCGCCGACACGCGCCCGCTCC 59
DB 562 GGCTCCACCGCGCGCCGCGACGCCAGGTGTCACTCGCGCCGACACCGCGCGCC 620

Search completed: January 16, 2005, 09:30:24
Job time : 184.4 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2005, 19:50:13 ; Search time 35.2 Seconds
(without alignments)
1211.572 Million cell updates/sec

Title: US-10-057-136-5

Perfect score: 60
Sequence: 1 ggaaccacggcgccgctgc.....ccgacacggcgccgctccc 60

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:
1: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/prodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/prodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/prodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/prodata/1/ina/6CTUS_COMB.seq:*
6: /cgn2_6/prodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35	58.3	1721	4 US-10-029-517-3	Sequence 3, Appl
2	35	58.3	1804	4 US-10-029-517-17	Sequence 17, Appl
3	35	58.3	8186	4 US-10-029-517-19	Sequence 19, Appl
4	33.6	56.0	518	4 US-10-029-517-101	Sequence 101, App
5	33.6	56.0	3343	4 US-10-029-517-102	Sequence 102, App
6	33.4	55.7	572	4 US-10-029-517-102	Sequence 18, Appl
7	32.2	53.7	60	4 US-09-475-947A-246	Sequence 246, App
8	32	53.3	981	4 US-10-029-517-16	Sequence 16, Appl
9	30.8	51.3	6192	2 US-08-479-537A-1	Sequence 1, Appl
10	30.8	51.3	6192	3 US-09-083-116-1	Sequence 1, Appl
11	30.8	51.3	6192	3 US-09-134-916A-1	Sequence 1, Appl
12	30.8	51.3	6449	2 US-08-479-537A-4	Sequence 4, Appl
13	30.8	51.3	6449	3 US-09-083-116-4	Sequence 4, Appl
14	30.8	51.3	6449	3 US-09-134-916A-4	Sequence 4, Appl
15	29.2	48.7	1023	1 US-08-252-966B-16	Sequence 16, Appl
16	27.8	46.3	2574	4 US-09-780-045-3	Sequence 3, Appl
17	27.2	44.5	1942	4 US-10-140-002-515	Sequence 515, App
18	26.4	44.0	1026	4 US-09-489-039A-1454	Sequence 1454, App
19	26.2	43.7	759	4 US-09-252-991A-5901	Sequence 5901, App
20	26.2	43.7	900	4 US-09-252-991A-5969	Sequence 5969, App
21	26.2	43.7	954	4 US-09-252-991A-5918	Sequence 5918, App
22	25.4	42.3	4403765	3 US-09-103-840A-2	Sequence 2, Appl
23	25.4	42.3	4403765	3 US-09-103-840A-2	Sequence 2, Appl
24	25.4	42.3	4411529	3 US-09-103-840A-1	Sequence 1, Appl
25	25.4	42.3	4411529	3 US-09-103-840A-1	Sequence 1, Appl
26	25.2	42.0	1029	4 US-09-252-991A-15460	Sequence 15460, A
27	25.2	42.0	1098	4 US-09-252-991A-15400	Sequence 15400, A

c	28	25.2	42.0	1488	4 US-09-252-991A-15378	Sequence 15378, A
	29	25.2	42.0	1842	4 US-09-252-991A-15508	Sequence 15508, A
	30	25.2	42.0	2877	3 US-09-235-103-1	Sequence 1, Appl
	31	25	41.7	1824	4 US-09-799-451-301	Sequence 301, App
	32	24.8	41.3	357	4 US-09-252-991A-441	Sequence 441, App
	33	24.8	41.3	1445	4 US-09-614-912-77	Sequence 77, Appl
	34	24.8	41.3	1446	4 US-09-252-991A-184	Sequence 184, Appl
	35	24.8	41.3	1488	4 US-09-252-991A-421	Sequence 421, App
	36	24.8	41.3	2825	4 US-09-196-390-5	Sequence 5, Appl
	37	24.8	41.3	2825	4 US-09-952-677-5	Sequence 5, Appl
	38	24.8	41.3	5993	3 US-09-383-630-1	Sequence 1, Appl
	39	24.8	41.3	5993	3 US-09-383-630-2	Sequence 2, Appl
	40	24.8	41.3	43280	2 US-08-804-227C-1	Sequence 1, Appl
	41	24.6	41.0	645	2 US-08-403-852D-9	Sequence 9, Appl
	42	24.6	41.0	645	3 US-08-510-646B-9	Sequence 9, Appl
	43	24.6	41.0	645	3 US-09-231-818-9	Sequence 9, Appl
	44	24.6	41.0	645	3 US-09-635-359B-9	Sequence 9, Appl
	45	24.6	41.0	2220	3 US-08-765-907A-14	Sequence 14, Appl

ALIGNMENTS

```
RESULT 1
US-10-029-517-3
; Sequence 3, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 3
; LENGTH: 1721
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (58) ... (1605)
US-10-029-517-3

Query Match      58.3%; Score 35; DB 4; Length 1721;
Best Local Similarity 74.6%; Pred. No. 0.041;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY      1  GGATCACCAGCGCGCGCTGCGACGAGTGCAGCGCCGACACGCGCCCGCTCC 59
Db      562  GGCTCACCAGCGCGCGCGACGACGAGTGCAGCTCGCGCCGACACGCGCCCGCTCC 620

RESULT 2
US-10-029-517-17
; Sequence 17, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 17
; LENGTH: 1804
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (73) ... (1500)
US-10-029-517-17
```

```
Query Match          58.3%; Score 35; DB 4; Length 1804;
Best Local Similarity 74.6%; Pred. No. 0.041;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY      1  GGATCCACCGCGCGCTGTGCGCACGAGTACGTGGCGCGCCGACACAGCGCGCTCC 59
Db      457 GGCTCCACCGCGCGCCCGCCAGCCACGATGTCACTCGGCGCCGACACAGCGCGCC 515

RESULT 3
US-10-029-517-19
; Sequence 19, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Susan J. Myers
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 19
; LENGTH: 8186
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 6899
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 7155
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 7184
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 7957
; OTHER INFORMATION: unknown
; NAME/KEY: intron
; LOCATION: (2997)...(3498)
; OTHER INFORMATION: intron 1
; NAME/KEY: intron:exon junction
; LOCATION: (3498)...(3499)
; OTHER INFORMATION: intron 1:exon 2
; NAME/KEY: exon
; LOCATION: (3508)...(3599)
; OTHER INFORMATION: exon 2d
; NAME/KEY: exon:intron junction
; LOCATION: (3982)...(3983)
; OTHER INFORMATION: exon 2a:intron 2a
; NAME/KEY: intron:exon junction
; LOCATION: (4205)...(4206)
; OTHER INFORMATION: intron 2c:exon 3c
; NAME/KEY: intron:exon junction
; LOCATION: (4259)...(4260)
; OTHER INFORMATION: intron 2d:exon 3d
; NAME/KEY: exon
; LOCATION: (4260)...(4328)
; OTHER INFORMATION: exon 3d
; NAME/KEY: intron:exon junction
; LOCATION: (4632)...(4633)
; OTHER INFORMATION: intron 3:exon 4
; NAME/KEY: exon
; LOCATION: (4914)...(5035)
; OTHER INFORMATION: exon 5
; NAME/KEY: intron
; LOCATION: (5266)...(6293)
; OTHER INFORMATION: intron 6
US-10-029-517-19

Query Match          58.3%; Score 35; DB 4; Length 8186;
Best Local Similarity 74.6%; Pred. No. 0.039;

Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY      1  GGATCCACCGCGCGCTGTGCGCACGAGTACGTGGCGCGCCGACACAGCGCGCTCC 59
Db      3825 GGCTCCACCGCGCGCCCGCCAGCCACGATGTCACTCGGCGCCGACACAGCGCGCC 3883

RESULT 4
US-10-029-517-101
; Sequence 101, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Susan J. Myers
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 101
; LENGTH: 518
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-10-029-517-101

Query Match          56.0%; Score 33.6; DB 4; Length 518;
Best Local Similarity 75.0%; Pred. No. 0.11;
Matches 42; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY      1  GGATCCACCGCGCGCTGTGCGCACGAGTACGTGGCGCGCCGACACAGCGCGCTCC 56
Db      462 GGCTCCACCGCGCGCCCGCCAGCCACGATGTCACTCGGCGCCCGACACAGCGCGCC 517

RESULT 5
US-10-029-517-102
; Sequence 102, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Susan J. Myers
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 102
; LENGTH: 3343
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-10-029-517-102

Query Match          56.0%; Score 33.6; DB 4; Length 3343;
Best Local Similarity 75.0%; Pred. No. 0.1;
Matches 42; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY      1  GGATCCACCGCGCGCTGTGCGCACGAGTACGTGGCGCGCCGACACAGCGCGCTCC 56
Db      1728 GGCTCCACCGCGCGCCCGCCAGCCACGATGTGTCACTCGGCGCCGACACAGCGCGCC 1783

RESULT 6
US-10-029-517-18
; Sequence 18, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Susan J. Myers
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517
```

CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 107
SEQ ID NO 18
LENGTH: 572
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (67)...(572)
US-10-029-517-18

Query Match 55.7%; Score 33.4; DB 4; Length 572;
Best Local Similarity 72.9%; Pred. No. 0.13;
Matches 43; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1 GGATCACCAGCCCGCCCTGCGCAGAGTGAGCTCGCCCGCCGACACGCGCCCTCC 59
DB 478 GGCTCCACCGCCCGCCCGCCAGCCGAGGTGTACCTCGCCCGGACACACGCGCCCGCC 536

RESULT 7
US-09-475-947A-246
Sequence 246, Application US/09475947A
Patent No. 6472154
GENERAL INFORMATION:
APPLICANT: Garner, Harold R.
APPLICANT: Wren, Jonathan D.
APPLICANT: Minna, John D.
TITLE OF INVENTION: Polymorphic Repeats in Human Genes
FILE REFERENCE: UTS0667
CURRENT APPLICATION NUMBER: US/09/475,947A
CURRENT FILING DATE: 1999-12-31
NUMBER OF SEQ ID NOS: 346
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 246
LENGTH: 60
TYPE: DNA
ORGANISM: human
US-09-475-947A-246

Query Match 53.7%; Score 32.2; DB 4; Length 60;
Best Local Similarity 75.5%; Pred. No. 0.3;
Matches 40; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 GGATCACCAGCCCGCCCTGCGCAGAGTGAGCTCGCCCGCCGACACGCGCC 53
DB 7 GGCTCCACCGCCCGCCCGCCAGCCGAGGTGTACCTCGCCCGGACACACGAGCC 59

RESULT 8
US-10-029-517-16
Sequence 16, Application US/10029517
Patent No. 6716627
GENERAL INFORMATION:
APPLICANT: Kenneth W. Dobie
APPLICANT: Susan J. Myers
TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
FILE REFERENCE: RTS-0352
CURRENT APPLICATION NUMBER: US/10/029,517
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 107
SEQ ID NO 16
LENGTH: 981
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: exon:exon junction
LOCATION: (464)...(465)
OTHER INFORMATION: exon 3b:exon 4
US-10-029-517-16

Query Match 53.3%; Score 32; DB 4; Length 981;
Best Local Similarity 73.2%; Pred. No. 0.32;

Matches 41; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 GGATCACCAGCCCGCCCTGCGCAGAGTGAGCTCGCCCGCCGACACGCGCCCGCC 56
DB 21 GGCTCCACCGCCCGCCCGCCAGCCGAGGTGTACCTCGCCCGGACACACGCGCCCGCC 76

RESULT 9
US-08-479-537A-1
Sequence 1, Application US/08479537A
Patent No. 5861381
GENERAL INFORMATION:
APPLICANT: CHAMON, Pierre
APPLICANT: KIENY, Marie-Paule
APPLICANT: LATHE, Richard
APPLICANT: HAREVENI, Mara
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
TREATMENT OR PREVENTION OF A MALIGNANT TUMOR
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: BURNS, DOANE, SWECKER & MATTHIS, L.L.P.
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/479,537A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 90/13101
FILING DATE: 23-OCT-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/FR91/00835
FILING DATE: 23-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/039,320
FILING DATE: 04-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/403,576
FILING DATE: 14-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Teskin, Robin L.
REGISTRATION NUMBER: 35,030
REFERENCE/DOCKET NUMBER: 017753-025
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 6192 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 58..120
FEATURE:
NAME/KEY: repeat region
LOCATION: 439..5239
OTHER INFORMATION: /note="The nucleotides spanning
OTHER INFORMATION: 439-5239 constitute a repeated region wherein the repeat is a
OTHER INFORMATION: nucleotides and encodes 20 amino acids, 17 of which are fixed
FEATURE:
NAME/KEY: mat_peptide

```
LOCATION: 121..6166
FEATURE:
NAME/KEY: repeat_region
LOCATION: 457
OTHER INFORMATION: /note= "Nucleotide 457 is X1 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 487
OTHER INFORMATION: /note= "Nucleotide 487 is Y = NNN
OTHER INFORMATION: which is the codon for Thr or Asn wherein Thr = ACT, ACC, ACA
OTHER INFORMATION: or ACG; and Asn = AAT or AAC."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 496
OTHER INFORMATION: /note= "Nucleotide 496 is X2 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
US-08-479-537A-1

Query Match
Best Local Similarity 51.3%; Score 30.8; DB 2; Length 6192;
Pred. No. 0.67;
Matches 38; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

Qy 1 GGATCCACCGCGCGCTGGCGACGAGTGACGTGGCGCGCCGACACGCGCGCTCC 59
Db 442 GGCTCCACCGCGCGCCCGCCCGCCGCGGTGTCACTCGCGCCCGGACNNNAGCCGANNCC 500

RESULT 10
US-09-083-116-1
Sequence 1, Application US/09083116
GENERAL INFORMATION:
APPLICANT: CHAMBON, Pierre
APPLICANT: KIENY, Marie-Paule
APPLICANT: LATHE, Richard
APPLICANT: HAREUVENI, Mara
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
TITL OF INVENTION: TREATMENT OR PREVENTION OF A MALIGNANT TUMOR
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRES:
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/083,116
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/479,537
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/FR91/00835
FILING DATE: 23-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/039,320
FILING DATE: 04-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/403,576
FILING DATE: 14-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Teskin, Robin L.
REGISTRATION NUMBER: 35,030
```

```
REFERENCE/DOCKET NUMBER: 017753-025
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 6192 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 58..120
FEATURE:
NAME/KEY: repeat_region
LOCATION: 439..5239
OTHER INFORMATION: /note= "The nucleotides spanning
OTHER INFORMATION: 439-5239 constitute a repeated region wherein the repeat is 6
OTHER INFORMATION: nucleotides and encodes 20 amino acids, 17 of which are fixed
OTHER INFORMATION: The number of such repeats varies from 1 to 80."
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 121..6166
FEATURE:
NAME/KEY: repeat_region
LOCATION: 457
OTHER INFORMATION: /note= "Nucleotide 457 is X1 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 487
OTHER INFORMATION: /note= "Nucleotide 487 is Y = NNN
OTHER INFORMATION: which is the codon for Thr or Asn wherein Thr = ACT, ACC, ACA
OTHER INFORMATION: or ACG; and Asn = AAT or AAC."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 496
OTHER INFORMATION: /note= "Nucleotide 496 is X2 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
US-09-083-116-1

Query Match
Best Local Similarity 51.3%; Score 30.8; DB 3; Length 6192;
Pred. No. 0.67;
Matches 38; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

Qy 1 GGATCCACCGCGCGCGCTGGCGACGAGTGACGTGGCGCGCCGACACGCGCGCTCC 59
Db 442 GGCTCCACCGCGCGCCCGCCCGCCGCGGTGTCACTCGCGCCCGGACNNNAGCCGANNCC 500

RESULT 11
US-09-134-916A-1
Sequence 1, Application US/09134916A
Patent No. 6328956
GENERAL INFORMATION:
APPLICANT: CHAMBON, Pierre
APPLICANT: KIENY, Marie-Paule
APPLICANT: LATHE, Richard
APPLICANT: HAREUVENI, Mara
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
TITL OF INVENTION: TREATMENT OR PREVENTION OF A MALIGNANT TUMOR
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRES:
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
```

```

1 MEDIUM TYPE: Floppy disk
2 COMPUTER: IBM PC compatible
3 OPERATING SYSTEM: PC-DOS/MS-DOS
4 SOFTWARE: PatentIn Release #1.0, Version #1.30
5 CURRENT APPLICATION DATA:
6 APPLICATION NUMBER: US/09/134,916A
7 FILING DATE:
8 CLASSIFICATION:
9 PRIOR APPLICATION DATA:
10 APPLICATION NUMBER: US/08/479,537
11 FILING DATE: 07-JUN-1995
12 APPLICATION NUMBER: FR 90/13101
13 FILING DATE: 23-OCT-1990
14 PRIOR APPLICATION DATA:
15 APPLICATION NUMBER: WO PCT/FR91/00835
16 FILING DATE: 23-OCT-1991
17 PRIOR APPLICATION DATA:
18 APPLICATION NUMBER: US 08/039,320
19 FILING DATE: 04-APR-1993
20 PRIOR APPLICATION DATA:
21 APPLICATION NUMBER: US 08/403,576
22 FILING DATE: 14-MAR-1995
23 ATTORNEY/AGENT INFORMATION:
24 NAME: Teskin, Robin L.
25 REGISTRATION NUMBER: 35,030
26 REFERENCE/DOCKET NUMBER: 017753-025
27 TELECOMMUNICATION INFORMATION:
28 TELEPHONE: (703) 836-6620
29 TELEFAX: (703) 836-2021
30 INFORMATION FOR SEQ ID NO: 1:
31 SEQUENCE CHARACTERISTICS:
32 LENGTH: 6192 base pairs
33 TYPE: nucleic acid
34 STRANDEDNESS: single
35 TOPOLOGY: linear
36 MOLECULE TYPE: DNA (genomic)
37 FEATURE:
38 NAME/KEY: sig_peptide
39 LOCATION: 58..120
40 FEATURE:
41 NAME/KEY: repeat_region
42 LOCATION: 439..5239
43 OTHER INFORMATION: /note= "The nucleotides spanning
44 OTHER INFORMATION: 439-5239 constitute a repeated region wherein the repeat is 6
45 OTHER INFORMATION: nucleotides and encodes 20 amino acids, 17 of which are fixed
46 OTHER INFORMATION: The number of such repeats varies from 1 to 80."
47 FEATURE:
48 NAME/KEY: mat_peptide
49 LOCATION: 121..6166
50 FEATURE:
51 NAME/KEY: repeat_region
52 LOCATION: 457
53 OTHER INFORMATION: /note= "Nucleotide 457 is X1 = NNN
54 OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCG, CCA
55 OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
56 FEATURE:
57 NAME/KEY: repeat_region
58 LOCATION: 487
59 OTHER INFORMATION: /note= "Nucleotide 487 is Y = NNN
60 OTHER INFORMATION: which is the codon for Thr or Asn wherein Thr = ACT, ACC, ACA, ACG
61 OTHER INFORMATION: or ACG; and Asn = AAT or AAC."
62 FEATURE:
63 NAME/KEY: repeat_region
64 LOCATION: 496
65 OTHER INFORMATION: /note= "Nucleotide 496 is X2 = NNN
66 OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCG, CCA
67 OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
68 US-09-134-916A-1
69 Query Match 51.3%; Score 30.8; DB 3; Length 6192;
70 Best Local Similarity 64.4%; Pred. No. 0.67;
71 Matches 38; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

```

```

1  GGATTCACCGGCGCCTGTGCGACGAGTACGTCGCGCGCCGACGACGCGCGCTCC 59
Db 442  GGCTTCACCGCCGCCNNNGCCACAGTGTCCTCGCGCCCGGACNNAGCGCANNCC 500

RESULT 12
US-08-479-537A-4
; Sequence 4, Application US/08479537A
; Patent No. 5861381
; GENERAL INFORMATION:
; APPLICANT: CHAMBON, Pierre
; APPLICANT: KIENY, Marie-Paule
; APPLICANT: LATHE, Richard
; APPLICANT: HAREVENT, Mara
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
; TITLE OF INVENTION: TREATMENT OR PREVENTION OF A MALIGNANT TUMOR
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/479,537A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 96/13101
; FILING DATE: 23-OCT-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/FR91/00835
; FILING DATE: 23-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/039,320
; FILING DATE: 04-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/403,576
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Teakin, Robin L.
; REGISTRATION NUMBER: 35,030
; REFERENCE/DOCKET NUMBER: 017753-025
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6449 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: sig peptide
; LOCATION: 58..120
; FEATURE:
; NAME/KEY: repeat region
; LOCATION: 439..5239
; OTHER INFORMATION: /note= "The nucleotides spanning
; OTHER INFORMATION: 439-5239 constitute a repeated region wherein the repeat is 6
; OTHER INFORMATION: nucleotides and encodes 20 amino acids, 17 of which are fixed
; OTHER INFORMATION: The number of such repeats varies from 1 to 80."
; FEATURE:
; NAME/KEY: mat peptide
; LOCATION: 121..5661
; FEATURE:

```

```

1 NAME/KEY: repeat_region
2 LOCATION: 457
3 OTHER INFORMATION: /note= "Nucleotide 457 is X1 = NNN
4 OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
5 OTHER INFORMATION: or CCG, and Ala = GCT, GCC, GCA, or GCG."
6 FEATURE:
7 NAME/KEY: repeat_region
8 LOCATION: 487
9 OTHER INFORMATION: /note= "Nucleotide 487 is Y = NNN
10 OTHER INFORMATION: which is the codon for Thr or Asn wherein Thr = ACT, ACC, ACA
11 OTHER INFORMATION: or ACG, and Asn = AAT or AAC."
12 FEATURE:
13 NAME/KEY: repeat_region
14 LOCATION: 496
15 OTHER INFORMATION: /note= "Nucleotide 496 is X2 = NNN
16 OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
17 OTHER INFORMATION: or CCG, and Ala = GCT, GCC, GCA, or GCG."
18 US-08-479-537A-4
19
20 Query March 51.3%; Score 30.8; DB 2; Length 6449;
21 Best Local Similarity 64.4%; Pred. No. 0.67; 21; Indels 0; Gaps 0;
22 Matches 38; Conservative 0; Mismatches 21;
23
24 Db 442 GGCTTCACCGCGCCCGCTGCGGAGTGAAGTGGGCGCCGACGCGCCCGCTCC 59
25 1 GGATTCACCGCGCGCCCGCTGCGGAGTGAAGTGGGCGCCGACGCGCCCGCTCC 59
26 442 GGCTTCACCGCGCCCGCTGCGGAGTGAAGTGGGCGCCGACGCGCCCGCTCC 500
27
28 RESULT 13
29 US-09-083-116-4
30 Sequence 4, Application US/09083116
31 Patent No. 6203795
32
33 GENERAL INFORMATION:
34 APPLICANT: CHAMON, Pierre
35 APPLICANT: KIENY, Marie-Paule
36 APPLICANT: LATHE, Richard
37 APPLICANT: HAREUVENI, Mara
38 TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
39 TITLE OF INVENTION: TREATMENT OR PREVENTION OF A MALIGNANT TUMOR
40 NUMBER OF SEQUENCES: 5
41 CORRESPONDENCE ADDRESSES:
42 ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
43 STREET: P. O. Box 1404
44 CITY: Alexandria
45 STATE: Virginia
46 COUNTRY: United States
47 ZIP: 22313-1404
48
49 COMPUTER READABLE FORM:
50 MEDIUM TYPE: Floppy disk
51 COMPUTER: IBM PC compatible
52 OPERATING SYSTEM: PC-DOS/MS-DOS
53 SOFTWARE: PatentIn Release #1.0, Version #1.30
54 CURRENT APPLICATION DATA:
55 APPLICATION NUMBER: US/09/083,116
56 FILING DATE:
57
58 CLASSIFICATION:
59 PRIOR APPLICATION DATA:
60 APPLICATION NUMBER: 08/479,537
61 FILING DATE:
62
63 PRIOR APPLICATION DATA:
64 APPLICATION NUMBER: WO PCT/FR91/00835
65 FILING DATE: 23-OCT-1991
66
67 PRIOR APPLICATION DATA:
68 APPLICATION NUMBER: US 08/039,320
69 FILING DATE: 04-APR-1993
70
71 PRIOR APPLICATION DATA:
72 APPLICATION NUMBER: US 08/403,576
73 FILING DATE: 14-MAR-1995
74
75 ATTORNEY/AGENT INFORMATION:
76 NAME: Teskin, Robin L.
77 REGISTRATION NUMBER: 35,030
78 REFERENCE/DOCKET NUMBER: 017753-025
79
80 TELECOMMUNICATION INFORMATION:

```

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

```

1 OPERATING SYSTEM: PC-DOS/MS-DOS
2 SOFTWARE: PatentIn Release #1.0, Version #1.30
3 CURRENT APPLICATION DATA:
4 APPLICATION NUMBER: US/09/134,916A
5 FILING DATE:
6 CLASSIFICATION:
7 PRIOR APPLICATION DATA:
8 APPLICATION NUMBER: US/08/479,537
9 FILING DATE: 07-JUN-1995
10 APPLICATION NUMBER: FR 90/13101
11 FILING DATE: 23-OCT-1990
12 PRIOR APPLICATION DATA:
13 APPLICATION NUMBER: WO PCT/FR91/00835
14 FILING DATE: 23-OCT-1991
15 PRIOR APPLICATION DATA:
16 APPLICATION NUMBER: US 08/039,320
17 FILING DATE: 04-APR-1993
18 PRIOR APPLICATION DATA:
19 APPLICATION NUMBER: US 08/403,576
20 FILING DATE: 14-MAR-1995
21 ATTORNEY/AGENT INFORMATION:
22 NAME: Teskin, Robin L.
23 REGISTRATION NUMBER: 35,030
24 REFERENCE/DOCKET NUMBER: 017753-025
25 TELECOMMUNICATION INFORMATION:
26 TELEPHONE: (703) 836-6620
27 TELEFAX: (703) 836-2021
28 INFORMATION FOR SEQ ID NO: 4:
29 SEQUENCE CHARACTERISTICS:
30 LENGTH: 6449 base pairs
31 TYPE: nucleic acid
32 STRANDEDNESS: single
33 TOPOLOGY: linear
34 MOLECULE TYPE: DNA (genomic)
35 FEATURE:
36 NAME/KEY: sig_peptide
37 LOCATION: 58..120
38 FEATURE:
39 NAME/KEY: repeat_region
40 LOCATION: 439..5239
41 OTHER INFORMATION: /note= "The nucleotides spanning
42 OTHER INFORMATION: 439-5239 constitute a repeated region wherein the repeat is 6
43 OTHER INFORMATION: nucleotides and encodes 20 amino acids, 17 of which are fixed
44 OTHER INFORMATION: The number of such repeats varies from 1 to 80."
45 FEATURE:
46 NAME/KEY: mat_peptide
47 LOCATION: 121..5661
48 FEATURE:
49 NAME/KEY: repeat_region
50 LOCATION: 457
51 OTHER INFORMATION: /note= "Nucleotide 457 is X1 = NNN
52 OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
53 OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
54 FEATURE:
55 NAME/KEY: repeat_region
56 LOCATION: 487
57 OTHER INFORMATION: /note= "Nucleotide 487 is Y = NNN
58 OTHER INFORMATION: which is the codon for Thr or Asn wherein Thr = ACT, ACC, ACG
59 OTHER INFORMATION: or ACG; and Asn = AAT or AAC."
60 FEATURE:
61 NAME/KEY: repeat_region
62 LOCATION: 496
63 OTHER INFORMATION: /note= "Nucleotide 496 is X2 = NNN
64 OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
65 OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
66 US-09-134-916A-4

```

Query Match	51.3%	Score 30.8;	DB 3;	Length 6449;
Best Local Similarity	64.4%;	Pred. No. 0.67;		
Matches 38;	Conservative 0;	Mismatches 21;	Indels 0;	Gaps 0;
1 GATTCACCGGCGCGCCGCGGAGTACGTCGCGGCCGACACGCGCCCGCTCC 59				

Db 442 GGCTCCACCGCCCCNNNGCCACGGTGTACCTCGGCCCGGACNNNAGGCCGNNNCC 500

RESULT 15
 US-08-252-966B-16/c
 Sequence 16, Application US/08252966B
 Patent No. 5624818
 GENERAL INFORMATION:
 APPLICANT: Eisenman, Robert N.
 APPLICANT: Hurlin, Peter J.
 APPLICANT: Ayer, Donald E.
 TITLE OF INVENTION: Regulatory Proteins that Dimerize with
 TITLE OF INVENTION: Mad or Max
 NUMBER OF SEQUENCES: 19
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Christensen, O'Connor, Johnson, and KindnessPLLC
 STREET: 1420 Fifth Ave., Suite 2800
 CITY: Seattle
 STATE: Washington
 COUNTRY: USA
 ZIP: 98101-2347
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/252,966B
 FILING DATE: 01-JUN-1994
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Shelton, Dennis K.
 REGISTRATION NUMBER: 26,397.
 REFERENCE/DOCKET NUMBER: FHC117694
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (206) 682-8100
 TELEFAX: (206) 224-0779
 INFORMATION FOR SEQ ID NO: 16:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1023 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA to mRNA
 DESCRIPTION: clone 20; see Figure 27
 HYPOTHETICAL: NO
 ORIGINAL SOURCE:
 ORGANISM: Mus musculus

Qy	2	GATTCACCGCGCCCGCTTCGCGACGAGTGACGTGCGCGCCGACACGCGCCCGCTTC	59
Db	67	GCTCATCTCTTCGCGCGCGCGACGCGAGACGCGCGCGCGCGACATCTCCCGCGCGCTTC	10
Query Match 48.7%; Score 29.2; DB 1; Length 1023;			
Best Local Similarity 69.0%; Pred. No. 2.1;			
Matches 40; Conservative 0; Mismatches 18; Indels 0; Gaps 0			

Search completed: January 16, 2005, 03:01:26
Job time : 45.2 secs

```

Oy      2  GAATCCACCGCGCGCGCTGGCGACAGGAGTGAAGTGGCGCGCCCGGACACGCGCGCGCTCC  59
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      67  GCTTCATCTGCGCGCGCGCGCGACGAGACGCGCGCGCGCGCGCTGCCCCGCGCGCTCC  10

```

This Page Blank (uspto)

OY 1 GGATCCACCGCGCGCTGCGCAGAGTGAAGTCGGCGCCCGACACGCGCCGCTCC 59
 Db 316 GGCTCCACCGCCCCCGACCCAGCGTGTCACTCGGCGCCCGAGACACGAGCGGCCCC 258
 RESULT 2
 LOCUS BM791359
 DEFINITION K-EST0071342 S21SNUS20 Homo sapiens cDNA clone S21SNUS20-14-A06 5',
 mRNA sequence.
 ACCESSION BM791359
 VERSION BM791359.1 GI:19139591
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 604)
 Oh,K.J., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
 Kim,Y.S., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
 21C Frontier Korean EST Project 2001
 TITLE Unpublished (2002)
 JOURNAL Contact: Kim YS
 COMMENT Genome Research Center
 Korea Research Institute of Bioscience & Biotechnology
 52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
 Tel: +82-42-860-4470
 Fax: +82-42-860-4409
 Email: yongseung@mail.krdb.re.kr
 Plate: 14 row: A column: 06
 High quality sequence stop: 604.
 Location/Qualifiers
 1..604
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="S21SNUS20-14-A06"
 /sex="F"
 /tissue_type="Stomach"
 /cell_type="Floating aggregates"
 /cell_line="SNU-520"
 /lab_host="SNU-520"
 /lab_host="Top10F"
 /clone_id="S21SNUS20"
 /note="Organ: Stomach; Vector: pTZ18RP1; Site 1: EcoRI;
 Site 2: NotI; The poly (A) + RNA was dephosphorylated with
 bacterial alkaline phosphatase (BAP) and then deapped
 intact mRNA was ligated with DNA-RNA linker including EcoR
 I site by treatment of T4 RNA ligase and the first strand
 cDNA was synthesized from oligo dt-selected mRNA by
 priming with dt-tailed vector. The dt-tailed vector was
 adjusted to have about 60nt. The cDNA vector was
 circularized with E. coli DNA ligase after digestion of
 EcoRI which site is also included in vector. An RNA strand
 converted to a DNA strand by Okayama-Berg method. The
 obtained cDNA vectors were used for transformation of
 competent cells E. coli Top10F by electroporation method.
 The cDNA libraries constructed by this method are
 full-length enriched cDNA library."
 ORIGIN
 Query Match 58.3%; Score 35; DB 4; Length 604;
 Best Local Similarity 74.6%; Pred. No. 2.5;
 Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
 OY 1 GGATCCACCGCGCGCTGCGCAGAGTGAAGTCGGCGCCCGACACGCGCCGCTCC 59
 Db 28 GGCTCCACCGACCCCGACCGAGTGTCACTCGGCGCCCGAGACACGAGCGGCCCC 86
 RESULT 3
 BU542454

LOCUS BU542454 877 bp mRNA linear EST 13-SEP-2002
 DEFINITION AGENCOURT 10322173 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6574322
 5', mRNA sequence.
 ACCESSION BU542454
 VERSION BU542454.1 GI:22852937
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 877)
 NIH-MGC http://mhc.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-r@mail.nih.gov
 Tissue Procurement: DCTD/DTP
 cDNA Library Preparation: Rubin Laboratory
 DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: L1CM2769 row: h column: 02
 High quality sequence stop: 760.
 Location/Qualifiers
 1..877
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6574322"
 /tissue_type="carcinoma, cell line"
 /lab_host="DH10B (pnage-resistant)"
 /clone_id="NIH_MGC_40"
 /note="Organ: prostate; Vector: pOT87; Site 1: XhoI;
 Site 2: EcoRI; cDNA made by oligo-dt priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCACGAG(G). Library constructed by
 Ling Hong in the Laboratory of Gerald W. Rubin (University
 of California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC library."
 ORIGIN
 Query Match 58.3%; Score 35; DB 5; Length 877;
 Best Local Similarity 74.6%; Pred. No. 2.4;
 Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
 OY 1 GGATCCACCGCGCGCTGCGCAGAGTGAAGTCGGCGCCCGACACGCGCCGCTCC 59
 Db 35 GGCTCCACCGCGCCCCCGACCGAGTGTCACTCGGCGCCCGAGACACGAGCGGCCCC 93
 RESULT 4
 CA489836 959 bp mRNA linear EST 14-NOV-2002
 LOCUS AGENCOURT 10810668 MAPCL Homo sapiens cDNA clone IMAGE:6722324 5',
 mRNA sequence.
 ACCESSION CA489836
 VERSION CA489836.1 GI:24952627
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 959)
 NIH-MGC http://mhc.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-r@mail.nih.gov
 Tissue Procurement: Krici A. Eglund, Ira Pastan
 cDNA Library Preparation: Invitrogen Corp

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM14284 row: n column: 20
High quality sequence start: 31
High quality sequence stop: 446.
Location/Qualifiers

FEATURES

source

```
1.959
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5722324"
/cell_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231,
hTERT-HME1, LNCaP"
/lab_host="EMD108"
/clone_lib="MAPCL"
/note="Vector: PCMV-SPORT6; Site 1: EcoRV; Site 2: Not I;
Subtracted with brain, liver, lung, kidney and muscle.
directionally cloned. Priming method: oligo-dT. Average
insert size: 1800 bp. Library amplification: 26,000 fold.
Kristi A. Eglund, James J. Vincent, Robert Strausberg,
Bungkok Lee & Ira Pastan: Discovery of new breast
cancer genes encoding membrane and secreted proteins.
Manuscript submitted."
```

ORIGIN

Query Match 58.3%; Score 35; DB 6; Length 959;
Best Local Similarity 74.6%; Pred. No. 2.4;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 1 GGATCACCAGCGCCGCTGCGACGAGTGAAGTGGCGCCGACACGCGCCGCTCC 59
Db 349 GGCTCACCAGCGCCCGCCGACGAGTGAAGTGGCGCCGACACGCGCGCCG 407

RESULT 5
LOCUS BUI48487 1113 bp mRNA linear EST 03-SEP-2002
DEFINITION AGENCOURT_8670479 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6380642
ACCESSION BUI48487
VERSION BUI48487.1 GI:22662019
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 1113)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov
Tissue Procurement: DCTD/DRP
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM2369 row: j column: 03
High quality sequence stop: 235.
Location/Qualifiers

FEATURES

source

```
1.1113
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6380642"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
```

/note="Organ: prostate; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN

Query Match 58.3%; Score 35; DB 5; Length 1113;
Best Local Similarity 74.6%; Pred. No. 2.4;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 1 GGATCACCAGCGCCGCTGCGACGAGTGAAGTGGCGCCGACACGCGCCGCTCC 59
Db 90 GGCTCACCAGCGCCCGCCGACGAGTGAAGTGGCGCCGACACGCGCGCCG 148

RESULT 6
LOCUS BUS42996 1130 bp mRNA linear EST 13-SEP-2002
DEFINITION AGENCOURT_10338707 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6574979
ACCESSION BUS42996
VERSION BUS42996.1 GI:22853479
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 1130)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov
Tissue Procurement: DCTD/DRP
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM2771 row: c column: 11
High quality sequence start: 27.
High quality sequence stop: 246.
Location/Qualifiers

FEATURES

source

```
1.1130
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6574979"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site 1: XhoI;  
Site 2: EcoRI; cDNA made by oligo-dT priming.  
directionally cloned into EcoRI/XhoI sites using the  
following 5' adaptor: GGCACGAG(G). Library constructed by  
ling Hong in the laboratory of Gerald M. Rubin (University  
of California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH_MGC Library."
```

ORIGIN

Query Match 58.3%; Score 35; DB 5; Length 1130;
Best Local Similarity 74.6%; Pred. No. 2.4;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 1 GGATCACCAGCGCCGCTGCGACGAGTGAAGTGGCGCCGACACGCGCCGCTCC 59
Db 69 GGCTCACCAGCGCCCGCCGACGAGTGAAGTGGCGCCGACACGCGCGCCG 127

RESULT 7
 LOCUS BQ936898 1234 bp mRNA linear EST 21-AUG-2002
 DEFINITION AGNCOURT 8919110 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6383855
 5', mRNA sequence.
 ACCESSION BQ936898
 VERSION BQ936898.1 GI:22352281
 EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 1234)
 NIH-MGC http://mgc.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: DCTD/DRP
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLCM2577 row: c column: 24
 High quality sequence stop: 245.
 Location/Qualifiers
 1..1234
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6383855"
 /tissue_type="carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_1ib="NIH_MGC_40"
 /note="Organ: prostate; Vector: pOT87; Site 1: XhoI;
 Site 2: EcoRI; CDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCAAGAG(G). Library constructed by
 Ling Hong in the laboratory of Gerald W. Rubin (University
 of California, Berkeley) using ZAP-CDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC library."

ORIGIN
 Query Match 58.3%; Score 35; DB 5; Length 1234;
 Best Local Similarity 74.6%; Pred. No. 2.3;
 Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 GGATCCACCGCGCGCTGCGACGAGTGCAGCGCGCGCGCGCGCGCTCC 59
 |||
 Db 90 GGCTCACCG 148

RESULT 8
 LOCUS BQ943554 1268 bp mRNA linear EST 21-AUG-2002
 DEFINITION AGNCOURT 8777815 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6384308
 5', mRNA sequence.
 ACCESSION BQ943554
 VERSION BQ943554.1 GI:22359032
 EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 1268)
 NIH-MGC http://mgc.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)

COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: DCTD/DRP
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLCM2579 row: b column: 21
 High quality sequence stop: 177.
 Location/Qualifiers
 1..1268
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6384308"
 /tissue_type="carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_1ib="NIH_MGC_40"
 /note="Organ: prostate; Vector: pOT87; Site 1: XhoI;
 Site 2: EcoRI; CDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCAAGAG(G). Library constructed by
 Ling Hong in the laboratory of Gerald W. Rubin (University
 of California, Berkeley) using ZAP-CDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC library."

ORIGIN
 Query Match 58.3%; Score 35; DB 5; Length 1268;
 Best Local Similarity 74.6%; Pred. No. 2.3;
 Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 GGATCCACCGCGCGCGCTGCGACGAGTGCAGCGCGCGCGCGCGCGCTCC 59
 |||
 Db 90 GGCTCACCG 148

RESULT 9
 LOCUS BQ920055 1343 bp mRNA linear EST 20-AUG-2002
 DEFINITION AGNCOURT 10031674 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6481609
 5', mRNA sequence.
 ACCESSION BQ920055
 VERSION BQ920055.1 GI:22334753
 EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 1343)
 NIH-MGC http://mgc.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: DCTD/DRP
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLCM2663 row: e column: 02
 High quality sequence start: 56
 High quality sequence stop: 237.
 Location/Qualifiers
 1..1343
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6481609"

FEATURES
 source

/cissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_1lb="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCAACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC library."

ORIGIN

Query Match 58.3%; Score 35; DB 5; Length 1343;
Best Local Similarity 74.6%; Pred. No. 2.3;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

1 GGATCACCAGCGCGCGCTGCGACGAGTGAAGTCTCGCGCCGACACGCGCCGCTCC 59
132 GGCTCCACCGCCCGCCCGACGCGTGTCACTCGCGCCGACACGCGCGCCGCTCC 190

RESULT 10

BU152566 1349 bp mRNA linear EST 03-SEP-2002
LOCUS AGENCOURT 8670683 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6380559
DEFINITION 5' mRNA sequence.

ACCESSION BU152566
VERSION BU152566.1 GI:22666098
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1349)
AUTHORS NIH-MGC http://mhc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHCW2569 row: f column: 16
High quality sequence stop: 291.

FEATURES

source

1.1349
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6380559"
/cissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_1lb="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCAACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC library."

ORIGIN

Query Match 58.3%; Score 35; DB 5; Length 1349;
Best Local Similarity 74.6%; Pred. No. 2.3;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

1 GGATCACCAGCGCGCGCTGCGACGAGTGAAGTCTCGCGCCGACACGCGCCGCTCC 59

19 GGCTCCACCGCCCGCCCGACGCGTGTCACTCGCGCCGACACGCGCGCCGCTCC 77

RESULT 11

BU542790 1420 bp mRNA linear EST 13-SEP-2002
LOCUS AGENCOURT 10334841 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6574725
DEFINITION 5' mRNA sequence.

ACCESSION BU542790
VERSION BU542790.1 GI:22853273
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1420)
AUTHORS NIH-MGC http://mhc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHCW2770 row: h column: 21
High quality sequence stop: 288.

FEATURES

source

1.1420
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6574725"
/cissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_1lb="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCAACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC library."

ORIGIN

Query Match 58.3%; Score 35; DB 5; Length 1420;
Best Local Similarity 74.6%; Pred. No. 2.3;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

1 GGATCACCAGCGCGCGCTGCGACGAGTGAAGTCTCGCGCCGACACGCGCCGCTCC 59
19 GGCTCCACCGCCCGCCCGACGCGTGTCACTCGCGCCGACACGCGCGCCGCTCC 77

RESULT 12
CF135986 648 bp mRNA linear EST 09-SEP-2003
LOCUS UI-HF-BNO-amo-h-10-0-UI.r1 NIH_MGC_50 Homo sapiens cDNA clone
DEFINITION IMAGE:3091050 5' mRNA sequence.

ACCESSION CF135986
VERSION CF135986.1 GI:33251430
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 648)
AUTHORS Bonaldo, M.F., Lennon, G. and Soares, M.B.

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cd_clone="IMAGE:6382088"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_1ib="NH_MGC_40"
/notes="Organ: prostate; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University

```


This Page Blank (uspto)


```
XX SQ Sequence 60 BP; 7 A; 29 C; 19 G; 5 T; 0 U; 0 Other;
Query Match 100.0%; Score 60; DB 2; Length 60;
Best Local Similarity 100.0%; Pred. No. 1.1e-07;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 GGATCCACCGCGCGCTGCGCAGAGTACGTCGCGCGCCGACACGCGCGCGCTGCC 60
DB 1 GGATCCACCGCGCGCTGCGCAGAGTACGTCGCGCGCCGACACGCGCGCGCTGCC 60

RESULT 2
AAV48324
ID AAV48324 standard; DNA; 60 BP.
XX AC AAV48324;
XX DT 20-NOV-1998 (first entry)
XX DE Nucleotide sequence encoding MUC1 tandem repeat unit R9.
XX KW ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
XX KM tumour; tumour-associated antigen.
XX OS Homo sapiens.
XX PN MO9837095-A2.
XX PD 27-AUG-1998.
XX PF 24-FEB-1998; 98WC-US0003693.
XX PR 24-FEB-1997; 97US-0038253P.
XX PA (THER-) THERION BIOLOGICS CORP.
XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX PA (DAND) DANA FARBER CANCER INST INC.
XX PI Schlom J, Kantor J, Kufe D, Panicali D, Gritz L,
XX DR WPI; 1998-467492/40.
XX PT New recombinant pox virus for tumour therapy - comprises DNA encoding an
XX PT immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
XX PS Disclosure; Page 11; 42pp; English.
XX CC The MUC1 tandem repeat units AAV48317-V48325 were used to create an
XX CC immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
XX CC (RPV). The RPV was used in a pharmaceutical composition also containing
XX CC an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The
XX CC recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
XX CC does not undergo significant genetic deletion, thereby providing an
XX CC unexpectedly stable and immunogenic pox virus. They can be used to
XX CC prevent or treat tumours expressing MUC1 tumour-associated antigens
XX SQ Sequence 60 BP; 12 A; 26 C; 14 G; 8 T; 0 U; 0 Other;
Query Match 63.7%; Score 38.2; DB 2; Length 60;
Best Local Similarity 78.0%; Pred. No. 0.12;
Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
OY 1 GGATCCACCGCGCGCTGCGCAGAGTACGTCGCGCGCCGACACGCGCGCGCTGCC 59
DB 1 GGTTCAACGGCAGCTTCAGCACGAGTACGTCGACCCGACACCCGTCAGGCTCC 59

RESULT 3
ADK68635/C
ID ADK68635 standard; CDNA; 156 BP.
XX AC ADK68635;

XX DT 06-MAY-2004 (first entry)
XX DE HSP65-MUC1 antigen CTL epitope related CDNA #3.
XX KW Human breast cancer; Mycobacterium bovis; heat shock protein 65; BCG;
XX KM HSP65; cytotoxic T lymphocyte; CTL; MUC1; HSP65-MUC1 antigen CTL epitope;
XX KW ss.
XX OS Unidentified.
XX PN CN1368384-A.
XX PD 11-SEP-2002.
XX PF 08-FEB-2001; 2001CN-00102614.
XX PR 08-FEB-2001; 2001CN-00102614.
XX PA (DIWE-) DIWEIHUAYU BIO TECHNOLOGY CO LTD BEIJING.
XX PI Yu Y, Li H;
XX DR WPI; 2003-854662/80.
XX PT Genetically engineered vaccine of MUC-1 antigen for human breast cancer.
XX PS Example 2; Page 4 (Disclosure); 14pp; Chinese.
XX CC The invention relates to a method of preparation of a genetically
XX CC engineered vaccine for preventing and treating human breast cancer. The
XX CC method comprises fusing the coding gene of the Mycobacterium bovis heat
XX CC shock protein 65 (BCG HSP65) with the cytotoxic T lymphocyte (CTL)
XX CC epitope gene of MUC1 (antigen cell expressed by human breast cancer
XX CC cells) and the HSP65-MUC1 antigen CTL epitope is expressed in Escherichia
XX CC coli cells. This sequence represents DNA used in the method of the
XX SQ Sequence 156 BP; 37 A; 42 C; 58 G; 19 T; 0 U; 0 Other;
Query Match 61.0%; Score 36.6; DB 10; Length 156;
Best Local Similarity 76.3%; Pred. No. 0.31;
Matches 45; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
OY 1 GGATCCACCGCGCGCTGCGCAGAGTACGTCGCGCGCCGACACGCGCGCGCTGCC 59
DB 132 GGTTCAACGGCAGCTTCAGCACGAGTACCTGTCGCGGACACCGCTCGGCTCC 74

RESULT 4
ADP32633
ID ADP32633 standard; DNA; 1818 BP.
XX AC ADP32633;
XX DT 26-FEB-2004 (first entry)
XX DE Plasmid JNM656 7x VNR MUC-1 nucleotide sequence.
XX KW MUC-1 antigen; immune response; MUC-1; variable number of tandem repeat;
XX KM VNR; repeat unit; tumour; metastasis; cytostatic; vaccine; gene therapy;
XX KW gene; ds.
XX OS Synthetic.
XX PN WO2003100060-A2.
XX PD 04-DEC-2003.
XX PF 23-MAY-2003; 2003WO-EP005594.
XX PR 24-MAY-2002; 2002GB-00012046.
```

PA (GLAX) GLAXO GROUP LTD.
XX
PI Burden N, Ellis JH, Hamblin PA;
XX
DR WPI; 2004-042811/04.
XX
PT New nucleic acid molecule encoding a MUC-1 antigen, useful for preparing
XX
PT a composition for treating or preventing tumors or metastases.
XX
PS Example; Fig 6; 66pp; English.

The present invention describes a nucleic acid molecule which encodes a MUC-1 antigen. The nucleic acid is capable of raising an immune response in vivo, has reduced susceptibility to recombination than full-length MUC -1 and comprises between 1 and 15 variable number of tandem repeats (VNT) perfect repeat units. Also described: (1) a plasmid comprising the DNA molecule; (2) a protein encoded by the nucleic acid; (3) a pharmaceutical composition comprising the nucleic acid, plasmid or protein and an excipient, diluent or carrier; and (4) a method of treating or preventing tumors or metastases. A MUC1 antigen has cytostatic activity, and can be used in vaccines, and in gene therapy. The nucleic acid is useful for preparing a composition for treating or preventing tumors or metastases. The present sequence is used in the exemplification of the present invention.

SQ Sequence 1818 BP, 346 A; 703 C; 430 G; 339 T; 0 U; 0 Other;

Query Match 61.0%; Score 36.6; DB 12; Length 1818;
Best Local Similarity 76.3%; Pred. No. 0.25;
Matches 45; Conservative 0; Mismatches 14; Indels 0; Gaps 0

Oy 1 GGATCCACCGCGCGCTGCGCAGAGTGAAGTCGGCGCCCGACACGCGCCGCTCC 59
Db 583 GGCTCACCGCGCCCCCGACGCCAGCATGTGCACCTCGCGCCCGAACAACGCGCCGCC 641

RESULT 5
ADF32629
ID ID ADF32629 standard; DNA; 2135 BP.
XX
XX ADF32629;
XX
XX 26-FEB-2004 (first entry)
DT DT
XX
DE Plasmid JMW319 7x VNT MUC-1 nucleotide sequence.

MUC-1 antigen; immune response; MUC-1; variable number of tandem repeat;
KM VNT; repeat unit; tumour; metastasis; cytostatic; vaccine; gene therapy;
XX gene; de.
XX
XX Synthetic.
XX
XX WO2003100060-A2.
PN
PD 04-DEC-2003.
XX
PF 23-MAY-2003; 2003WO-EP005594.
XX
PR 24-MAY-2002; 2002GB-00012046.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI Burden N, Ellis JH, Hamblin PA;
XX
PI WPI; 2004-042811/04.
DR
XX
XX New nucleic acid molecule encoding a MUC-1 antigen, useful for preparing
PT a composition for treating or preventing tumors or metastases.
XX
XX Example; Fig 4; 66pp; English.

The present invention describes a nucleic acid molecule which encodes a MUC-1 antigen. The nucleic acid is capable of raising an immune response

CC	in vivo comprises reduced susceptibility to recombination than full-length MD
CC	-1 and comprises between 1 and 15 variable number of tandem repeats
CC	(VNR) perfect repeat units. Also described: (1) a plasmid comprising the
CC	DNA molecule; (2) a protein encoded by the nucleic acid; (3) a
CC	pharmaceutical composition comprising the nucleic acid, plasmid or
CC	protein and an excipient, diluent or carrier; and (4) a method of
CC	treating or preventing tumours or metastases. A MUC1 antigen has
CC	cytostatic activity, and can be used in vaccines, and in gene therapy.
CC	The nucleic acid is useful for preparing a composition for treating or
CC	preventing tumours or metastases. The present sequence is used in the
XX	embodiment of the present invention.
SQ	Sequence 2135 BP; 395 A; 812 C; 524 G; 404 T; 0 U; 0 Other;
OY	Query Match 61.0%; Score 36.6; DB 12; Length 2135;
Db	Best Local Similarity 76.3%; Pred. No. 0.24;
	Matches 45; Conservative 0; Mismatches 14; Indels 0; Gaps 0
	1 GGATCACCAGCGGCGCCCTGGACGAGACGTGCGCGCCGCAGACAGGCCGCTCC 59
	642 GACTTCACCGCCCCCGCCAGCCACGCTGTCACTCGGCCCGGACACAGCCGCCCC 700
RESULT 6	
ID	AD157667 standard; cDNA, 2255 BP.
XX	AD157667;
DT	22-APR-2004 (first entry)
DE	Human breast specific nucleic acid (BSNA) #28.
KW	Human; breast specific nucleic acid; BSNA; gene; ss; metastasis;
OS	breast cancer; cytostatic.
PN	Homo sapiens.
PD	WO2003106648-A2.
PF	24-DEC-2003.
PR	16-JUN-2003; 2003WO-US018934.
PA	14-JUN-2002; 2002US-0389327P.
PB	(DIAD-) DIADEXUS INC.
PI	Salceda S, Macina RA, Turner LR, Sun Y, Liu C;
DR	MPJ; 2004-082185/08.
PT	P-PsDB; AD157741.
PS	Novel isolated polypeptide comprising breast specific protein sequences,
PP	useful for diagnosing or monitoring presence and metastases of breast
PT	cancer in patient.
PS	Claim 1; SEQ ID NO 38; 370pp; English.
XX	
CC	The invention relates to human breast specific nucleic acids (BSNA) and
CC	the breast specific proteins (BSP) they encode. The nucleic acids are
CC	useful for determining the presence of a BSNA in a sample which involves
CC	contacting the sample with a BSNA under conditions in which the BSNA will
CC	selectively hybridise to a BSNA in the sample, and detecting the
CC	hybridisation. The nucleic acids are useful for determining the presence
CC	of a BSP in a sample which involves contacting the sample with suitable
CC	reagent under conditions in which the reagent will selectively interact
CC	with the BSP, and detecting the interaction of the reagent with a BSP in
CC	the sample. The nucleic acids and proteins are useful for diagnosing or
CC	monitoring the presence and metastases of breast cancer in a patient,
CC	which involves determining an amount of nucleic acid or protein and
CC	comparing the determined amount of nucleic acid or protein in the sample
CC	of the patient to the amount of a breast specific marker in a normal

CC	control, where a difference in the determined amount in the sample
CC	compared to the amount in the control is associated with the presence of
CC	breast cancer. The sequences are useful for treating a patient with
CC	breast cancer, involving administering a composition consisting of a BSN
CC	or a BSP to a patient, where the administration induces an immune
CC	response against the breast cancer cell expressing the BSN or BSP. This
CC	sequence represents a human BSN of the invention.
CC	
XX	Sequence 2255 BP, 442 A, 763 C, 585 G, 461 T, 0 U, 4 Other;
XX	
XX	Query Match
XX	Beat Local Similarity 59.7%; Score 35.8; DB 12; Length 2255;
XX	Matches 43; Conservative 2; Mismatches 14; Indels 0; Gaps
Db	
Oy	1 GGATCCACCGCGCCCTGCGCAGAGTACGTCGCGCCGACACGCGCCCGCTCC 59
Db	578 GGCTCCACCGCGCCCGCCGACGCGTGTCACTCGAGCCCGGACACGAGCGGCCCC 636
RESULT 7	
ID	AD157666 standard; cDNA; 4144 BP.
XX	
AC	AD157666;
XX	
DT	22-APR-2004 (first entry)
XX	
DE	Human breast specific nucleic acid (BSNA) #37.
XX	
KM	Human, breast specific nucleic acid; BSN; gene; ss; metastasis;
KW	breast cancer; cytostatic.
OS	Homo sapiens.
XX	
PN	WO2003106648-A2.
XX	
PD	24-DEC-2003.
XX	
PF	16-JUN-2003; 2003WO-US018934.
XX	
PR	14-JUN-2002; 2002US-0389327P.
XX	
PA	(DIAD)- DIADEXUS INC.
PI	Salceda S, Macina RA, Turner LR, Sun Y, Liu C;
DR	WPI; 2004-082185/08.
PT	P-PADB; AD157740.
XX	
PS	Novel isolated polypeptide comprising breast specific protein sequences,
PT	useful for diagnosing or monitoring presence and metastases of breast
XX	cancer in patient.
XX	
XX	Claim 1; SEQ ID NO 37; 370pp; English.
XX	
XX	The invention relates to human breast specific nucleic acids (BSNA) and
XX	the breast specific proteins (BSP) they encode. The nucleic acids are
XX	useful for determining the presence of a BSN in a sample which involves
XX	contacting the sample with a BSN under conditions in which the BSN will
XX	selectively hybridise to a BSN in the sample, and detecting the
XX	hybridisation. The nucleic acids are useful for determining the presence
XX	of a BSP in a sample which involves contacting the sample with suitable
XX	reagent under conditions in which the reagent will selectively interact
XX	with the BSP, and detecting the interaction of the reagent with a BSP in
XX	the sample. The nucleic acids and proteins are useful for diagnosing or
XX	monitoring the presence and metastases of breast cancer in a patient,
XX	which involves determining an amount of nucleic acid or protein and
XX	comparing the determined amount of nucleic acid or protein in the sample
XX	of the patient to the amount of a breast specific marker in a normal
XX	control, where a difference in the determined amount in the sample
XX	compared to the amount in the control is associated with the presence of
XX	breast cancer. The sequences are useful for treating a patient with
XX	breast cancer, involving administering a composition consisting of a BSN

Query	Best Local Similarity	Score	DB 2;	Length	Indels	Gaps
1 GATTCACCGCGCGCTGCGACGAGTGAAGTGGCGCCCGACACGCGCCCGCTCC	76.8%	58.7%	DB 2;	1320;	13;	0;
Matches 43;	Conservative 0;	Mismatches 14;	Indels 0;	Gaps 0;		

or a BSP to a patient, where the administration induces an immune response against the breast cancer cell expressing the BSNA or BSP. This sequence represents a human BSNA of the invention.

Sequence 4144 BP; 634 A; 1896 C; 1058 G; 552 T; 0 U; 4 Other;

Query Match 59.7%; Score 35.8; DB 12; Length 4144;
 Best Local Similarity 72.9%; Pred. No. 0.38;
 Matches 43; Conservative 2; Mismatches 14; Indels 0; Gaps 0;

1 GATTCACCGCGCGCTGCGACGAGTGAAGTGGCGCCCGACACGCGCCCGCTCC 59
 633 GGCTTCACCGCGCGCCSCAGCCACGAGTGAAGTGGCGCCCGACACGAGCGCGCC 691

RESULT 8
 AAT29190
 ID AAT29190 standard; cDNA to mRNA; 1320 BP.
 XX
 AC AAT29190;
 XX
 DT 26-JUL-1996 (first entry)
 XX
 DE Glycoprotein 39 3' fragment.
 XX
 KW Glycoprotein 39; gp39; lambda gtl1 cDNA library; gastric cancer;
 KM cell line KATO-III; tumour; immune abnormality; marker;
 XX inflammatory disease; ds.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..984
 FT /*tag= a
 FT polyA_signal 1267..1272
 FT /*tag= b
 FT polyA_signal 1293..1320
 FT /*tag= c
 XX
 PN JP07051065-A.
 XX
 PD 28-FEB-1995.
 XX
 PF 21-FEB-1992; 92JP-00035085.
 XX
 PR 21-FEB-1992; 92JP-00035085.
 XX
 PA (NIKO-) NIPPON KOTAI KENKYUSHO KK.
 PA (UYKA-) UNIV KAGOSHIMA.
 XX
 DR WPI; 1995-127356/17.
 DR P-PEDB; AAR96298.
 PT Glyco:protein 39 gene - used in the mass production of glyco:protein 39,
 PT for use as tumour and immune abnormality marker.
 PS Claim 2; Page 9-10; 10pp; Japanese.
 XX
 CC The sequences given in AAT29189-90 encode portions of glycoprotein 39.
 CC This sequence represents the 3' portion of the glycoprotein 39 gene and
 CC was isolated from lambda gtl1 cDNA library derived from the gastric
 CC cancer cell line KATO-III. Glycoprotein 39 is expected to be used as a
 CC tumour marker, an immune abnormality marker or a marker for various
 CC inflammatory diseases
 XX
 SQ Sequence 1320 BP; 288 A; 417 C; 317 G; 298 T; 0 U; 0 Other;

DB 1 GGCTCCACCGCCCCCGACGCGTGTACCTCGGCCCCGGACACAGCCCGC 56

RESULT 9
AAV48316
XX AAV48316 standard; cDNA; 60 BP.
XX
AC AAV48316;
XX
DT 20-NOV-1998 (first entry)
XX
DE Nucleotide sequence encoding MUC1 tandem repeat unit.
XX
KW ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
KM tumour; tumour-associated antigen.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH 1. 60
FT /tag= a
FT /product= "MUC1 tandem repeat unit"
FT
FT
FT
PN WO9837095-A2.
XX
XX 27-AUG-1998.
XX
XX 24-FEB-1998; 98WO-US003693.
XX
XX 24-FEB-1997; 97US-0038253P.
XX
XX (THER-) THERION BIOLOGICS CORP.
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA (DAND) DANA FARBER CANCER INST INC.
XX
XX Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;
PI MPI: 1998-467492/40.
DR P-PSDB; AAM77229.
XX
XX New recombinant pox virus for tumour therapy - comprises DNA encoding an
PT immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
PT
XX
XX Example 1; Page 20; 42pp; English.
XX
XX The MUC1 tandem repeat unit was used to create an immunogenic mini-MUC1
CC fragment for inclusion in a recombinant pox virus (RPV). The RPV was used
CC in a pharmaceutical composition also containing an immunomodulator to
CC generate MUC1 specific cytotoxic T-lymphocytes. The recombinant pox virus
CC therefore encodes an immunogenic MUC1 fragment that does not undergo
CC significant genetic deletion, thereby providing an unexpectedly stable
CC and immunogenic pox virus. They can be used to prevent or treat tumours
CC expressing MUC1 tumour-associated antigens
XX
SQ Sequence 60 BP; 7 A; 33 C; 16 G; 4 T; 0 U; 0 Other;

Query Match 58.3%; Score 35; DB 2; Length 60;
Best Local Similarity 74.6%; Pred. No. 0.93;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 GGATCCACCGCGCGCTGGCGACGAGTGACGTGCGCGCCCGACACGCGCCGCTCC 59
DB 1 GGCTCCACCGCCCCCGACGCGTGTACCTCGGCCCCGGACACAGCCCGC 59

RESULT 10
AAT29189
ID AAT29189 standard; cDNA to mRNA; 180 BP.
XX
XX AAT29189;
AC
XX
XX 25-JUL-1996 (first entry)
DT
XX

DE Glycoprotein 39 5' fragment.
XX
XX Glycoprotein 39; gp39; lambda gt11 cDNA library; gastric cancer;
KM cell line KATO-III; tumour; immune abnormality; marker;
KW inflammatory disease; ds.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH repeat_region 1..180
FT /tag= a
FT repeat_unit 1..60
FT /tag= b
XX
XX JP07051065-A.
XX
XX 28-FEB-1995.
PD
XX
XX 21-FEB-1992; 92JP-00035085.
PF
XX
XX 21-FEB-1992; 92JP-00035085.
PR
XX
XX (NIKO-) NIPPON KOTAI KENKYUSHO KK.
PA (UYKA-) UNIV KAGOSHIMA.
XX
XX MPI: 1995-127356/17.
DR P-PSDB; AAR6297.
XX
XX Glyco:protein 39 gene - used in the mass production of glyco:protein 39,
PT for use as tumour and immune abnormality marker.
PT
XX
XX Claim 2; Page 8; 10pp; Japanese.
PS
XX
XX The sequences given in AAT29189-90 encode portions of glycoprotein 39.
CC This sequence represents the 5' portion of the glycoprotein 39 gene and
CC was isolated from lambda gt11 cDNA library derived from the gastric
CC cancer cell line KATO-III. Glycoprotein 39 is expected to be used as a
CC tumour marker, an immune abnormality marker or a marker for various
CC inflammatory diseases
XX
SQ Sequence 180 BP; 21 A; 91 C; 56 G; 12 T; 0 U; 0 Other;

Query Match 58.3%; Score 35; DB 2; Length 180;
Best Local Similarity 74.6%; Pred. No. 0.84;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 GGATCCACCGCGCGCTGGCGACGAGTGACGTGCGCGCCCGACACGCGCCGCTCC 59
DB 121 GGCTCCACCGCGCGCGCGACGCGTGTACCTCGGCCCCGGACACAGCCCGC 179

RESULT 11
AAN90579/C
ID AAN90579 standard; cDNA; 309 BP.
XX
XX AAN90579;
AC
XX
XX 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 04-DEC-1989 (first entry)
DT
XX
XX PDF9.3 cDNA insert.
DE
XX
XX PDF9.3; human DF3 breast carcinoma-associated antigen epitope.
KM Human MCF-7 breast carcinoma cells; 'ATCC HTB22'.
OS
XX
XX WO8907107-A.
PN
XX
XX 10-AUG-1989.
PD
XX
XX 29-JAN-1988; 88US-00149831.
PF
XX

CC the sample. The nucleic acids and proteins are useful for diagnosing or
CC monitoring the presence and metastases of breast cancer in a patient,
CC which involves determining an amount of nucleic acid or protein in the sample
CC comparing the determined amount of nucleic acid or protein in the sample
CC of the patient to the amount of a breast specific marker in a normal
CC control, where a difference in the determined amount in the sample
CC compared to the amount in the control is associated with the presence of
CC breast cancer. The sequences are useful for treating a patient with
CC breast cancer, involving administering a composition consisting of a BSN
CC or a BSP to a patient, where the administration induces an immune
CC response against the breast cancer cell expressing the BSN or BSP. This
CC sequence represents a human BSN of the invention.

CC Sequence 1194 BP; 243 A; 438 C; 275 G; 238 T; 0 U; 0 Other;

Query Match 58.3%; Score 35; DB 12; Length 1194;
Best Local Similarity 74.6%; Pred. No. 0.71;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 GGATCCACCGCGCGCTGCGACGAGTGACGTGCGCGCCGACACGCGCCGCTCC 59
Db 596 GGCTCCACCGCGCGCGCTGCGACGAGTGACGTGACGTGCGCGCCGACACGCGCGCCG 654

RESULT 14

ID AD157693 standard; cDNA; 1378 BP.

AC AD157693;

DT 22-APR-2004 (first entry)

DE Human breast specific nucleic acid (BSNA) #64.

KW Human; breast specific nucleic acid; BSN; gene; ss; metastasis;
KM breast cancer; cytostatic.

OS Homo sapiens.

PN WO2003106648-A2.

XX 24-DEC-2003.

PF 16-JUN-2003; 2003WO-US018934.

PR 14-JUN-2002; 2002US-0389327P.

PA (DIAD-) DIADEXUS INC.

PI Salceda S, Macina RA, Turner LR, Sun Y, Liu C;

DR WPI; 2004-082185/08.
DR P-PSDB; AD157765.

PT Novel isolated polypeptide comprising breast specific protein sequences,
PT useful for diagnosing or monitoring presence and metastases of breast
PT cancer in patient.

PS Claim 1; SEQ ID NO 64; 370pp; English.

CC The invention relates to human breast specific nucleic acids (BSNA) and
CC the breast specific proteins (BSP) they encode. The nucleic acids are
CC useful for determining the presence of a BSN in a sample which involves
CC contacting the sample with a BSN under conditions in which the BSN will
CC selectively hybridize to a BSN in the sample, and detecting the
CC hybridization. The nucleic acids are useful for determining the presence
CC of a BSP in a sample which involves contacting the sample with suitable
CC reagent under conditions in which the reagent will selectively interact
CC with the BSP, and detecting the interaction of the reagent with a BSP in
CC the sample. The nucleic acids and proteins are useful for diagnosing or
CC monitoring the presence and metastases of breast cancer in a patient,
CC which involves determining an amount of nucleic acid or protein in the sample
CC comparing the determined amount of nucleic acid or protein in the sample

CC of the patient to the amount of a breast specific marker in a normal
CC control, where a difference in the determined amount in the sample
CC compared to the amount in the control is associated with the presence of
CC breast cancer. The sequences are useful for treating a patient with
CC breast cancer, involving administering a composition consisting of a BSN
CC or a BSP to a patient, where the administration induces an immune
CC response against the breast cancer cell expressing the BSN or BSP. This
CC sequence represents a human BSN of the invention.

CC Sequence 1378 BP; 295 A; 489 C; 316 G; 278 T; 0 U; 0 Other;

Query Match 58.3%; Score 35; DB 12; Length 1378;
Best Local Similarity 74.6%; Pred. No. 0.7;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 GGATCCACCGCGCGCTGCGACGAGTGACGTGCGCGCCGACACGCGCCGCTCC 59
Db 596 GGCTCCACCGCGCGCGCTGCGACGAGTGACGTGACGTGCGCGCCGACACGCGCGCCG 654

RESULT 15

ID ADO23180/c standard; RNA; 1424 BP.

AC ADO23180;

DT 12-AUG-2004 (first entry)

DE Antisense human MUC1 mucin glycoprotein RNA (coding sequence) SeqID 75.

KW human; MUC1; mucin glycoprotein; cancer; chemotherapeutic; MUC1/BCD; ss;
KM antisense.

OS Homo sapiens.

PN WO2004044160-A2.

XX 27-MAY-2004.

PF 12-NOV-2003; 2003WO-US035848.

PR 13-NOV-2002; 2002US-00293391.

PR 29-MAY-2003; 2003US-00447839.

PA (DAND) DANA FARBER CANCER INST INC.

PI (ILEX-) ILEX PROD INC.

PI Kufe DW, Kharbanda S, Weltman SD;

DR WPI; 2004-420304/39.

PT Double-stranded RNA complex useful for inhibiting proliferation of cancer
PT cell expressing MUC1 mucin glycoprotein, comprises first and second RNA
PT sequences.

PS Disclosure; SEQ ID NO 75; 112pp; English.

CC This invention relates to novel modulators of the human MUC1 mucin
CC glycoprotein for use in cancer therapeutics, where MUC1 is a protein that
CC acts to inhibit the apoptotic response to genotoxic stress caused by
CC chemotherapeutic agents. In particular, it refers to modulators of the
CC MUC1 extracellular domain (MUC1/BCD). The method refers to using double-
CC stranded RNA complexes as MUC1 interference RNA compositions such that
CC MUC1 expression is inhibited, which in turn inhibits cancer cell
CC proliferation. The present invention describes screening assays to
CC identify compounds that inhibit the binding of various MUC1 ligands such
CC as neutrophil 2, as well as agonists, antagonists and antibodies thereof.
CC Furthermore, it provides MUC1 antisense, small interfering RNA (siRNA)
CC and small molecules in combination with chemotherapeutic agents that are
CC useful in the field of cancer therapy. This polynucleotide sequence is
CC the antisense human MUC1 RNA of the invention.

Query Match 58.3%; Score 35; DB 12; Length 1424;
Best Local Similarity 74.6%; Pred. No. 0.7;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 1 GGATCCACCGGCGCCTGTGGCAGGAGTGAACGTGGGCGCCGACACGCGCCCGCTCC 59
Db 1041 GGCTCCACCGGCGCCCGCAGCCAGGTGTCACTCTGGGCGGACACGAGCGCGCC 983

Search completed: January 15, 2005, 20:36:07
Job time : 170.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2005, 22:37:00 ; Search time 183.4 Seconds
(without alignments)
1879.789 Million cell updates/sec

Title: US-10-057-136-4

Perfect score: 60

Sequence: 1 GGCAGTACTGCACCGCACCGC.....CTGATACAAGACTGCACCT 60

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 4300275 seqs, 2872944193 residues

Total number of hits satisfying chosen parameters: 8600550

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications NA.*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
21: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	60	100.0	60	US-10-057-136-4	Sequence 4, Appli
2	60	100.0	2297	US-10-406-317-41	Sequence 41, Appli
3	37.6	62.7	60	US-10-057-136-12	Sequence 12, Appli
4	35.4	59.0	60	US-10-057-136-10	Sequence 10, Appli
5	33.4	55.7	60	US-10-057-136-8	Sequence 8, Appli
6	33.4	55.7	60	US-10-057-136-9	Sequence 9, Appli
7	33.2	55.3	60	US-10-057-136-7	Sequence 7, Appli
8	29	48.3	60	US-10-057-136-11	Sequence 11, Appli
9	28.6	47.7	1230	US-10-369-493-42059	Sequence 42059, A
10	27.4	45.7	780	US-10-437-963-15324	Sequence 15324, A
11	27.4	45.7	60	US-10-057-136-6	Sequence 6, Appli
12	27.4	45.7	1682	US-10-479-638-10	Sequence 10, Appli

13	27	45.0	60	US-10-057-136-2	Sequence 2, Appli
14	27	45.0	78	US-10-057-136-13	Sequence 13, Appli
15	27	45.0	1424	US-10-447-839A-75	Sequence 75, Appli
16	27	45.0	1428	US-10-447-839A-20	Sequence 20, Appli
17	27	45.0	1527	US-10-057-136-19	Sequence 19, Appli
18	27	45.0	1799	US-10-447-839A-19	Sequence 19, Appli
19	27	45.0	1804	US-09-964-824A-573	Sequence 573, Appli
20	27	45.0	1804	US-10-029-517-17	Sequence 17, Appli
21	27	45.0	1804	US-10-717-597-10	Sequence 30, Appli
22	27	45.0	1804	US-10-775-920-84	Sequence 84, Appli
23	27	45.0	1823	US-10-101-510-339	Sequence 339, Appli
24	27	45.0	4139	US-09-964-824A-105	Sequence 105, Appli
25	27	45.0	4139	US-09-964-824A-578	Sequence 578, Appli
26	27	45.0	4139	US-09-864-864-334	Sequence 334, Appli
27	27	45.0	4139	US-09-880-107-2121	Sequence 2121, Appli
28	27	45.0	4139	US-09-968-007A-751	Sequence 751, Appli
29	27	45.0	4139	US-10-171-311-157	Sequence 157, Appli
30	27	45.0	4139	US-10-177-293-310	Sequence 310, Appli
31	27	45.0	4139	US-10-440-464-155	Sequence 155, Appli
32	27	45.0	4139	US-10-734-564-53	Sequence 53, Appli
33	27	45.0	4139	US-10-775-920-80	Sequence 80, Appli
34	27	45.0	4139	US-10-775-920-85	Sequence 85, Appli
35	27	45.0	8181	US-10-447-839A-18	Sequence 18, Appli
36	27	45.0	8186	US-10-247-703-23	Sequence 23, Appli
37	27	45.0	8186	US-10-029-517-19	Sequence 19, Appli
38	26.4	44.0	389	US-09-918-995-17681	Sequence 37681, A
39	26.4	44.0	922	US-10-424-599-61938	Sequence 61938, A
40	26.4	44.0	3722	US-10-305-720-1129	Sequence 1129, A
41	26.2	43.7	596	US-10-425-115-28374	Sequence 28374, A
42	26	43.3	1194	US-10-156-761-5206	Sequence 5206, A
43	26	43.3	9025608	US-10-156-761-1	Sequence 1, Appli
44	25.6	42.7	60	US-10-057-136-14	Sequence 14, Appli
45	25.6	42.7	518	US-10-247-703-38	Sequence 38, Appli

ALIGNMENTS

RESULT 1
US-10-057-136-4
Sequence 4, Application US/10057136
Publication No. US20030021770A1
GENERAL INFORMATION:
APPLICANT: SCHLON, JEFFREY
APPLICANT: KANTOR, JUDITH
APPLICANT: KUFE, DONALD
APPLICANT: PANICALI, DENNIS
TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
TITLE OF INVENTION: TUMOR-ASSOCIATED ANTIGEN
FILE REFERENCE: 700953/47113C
CURRENT APPLICATION NUMBER: US/10/057,136
CURRENT FILING DATE: 2002-01-25
PRIOR APPLICATION NUMBER: 09/366,670
PRIOR FILING DATE: 1999-08-03
PRIOR APPLICATION NUMBER: PCT/US98/03693
PRIOR FILING DATE: 1998-02-24
PRIOR APPLICATION NUMBER: 60/038,253
PRIOR FILING DATE: 1997-02-24
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 60
TYPE: DNA
ORGANISM: Homo sapiens
US-10-057-136-4
Query Match 100.0%; Score 60; DB 14; Length 60;
Best Local Similarity 100.0%; Pred. No. 3e-11;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 GGCAGTACTGCACCGCACCGCATGCTATCATGACGACCTGATACAGACTGCACCT 60
|||||

Db 1 GGCAGTACTGCACACCGCGACATGGCGTAATCATGACGACTGTATACAGACTTGACACT 60

RESULT 2

US-10-406-317-41
; Sequence 41, Application US/10406317
; Publication No. US20040019195A1
; GENERAL INFORMATION:
; APPLICANT: Schlom, Jeffrey;
; APPLICANT: Hodge, James;
; APPLICANT: Panicali, Dennis
; TITLE OF INVENTION: A recombinant vector expressing multiple constitutulatory
; FILE REFERENCE: 38163-0189
; CURRENT APPLICATION NUMBER: US/10/406,317
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: US/09/856,988
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: PCT/US99/26866
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/111,582
; PRIOR FILING DATE: 1998-12-09
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41
; LENGTH: 2297
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: VECTOR
US-10-406-317-41

Query Match 100.0%; Score 60; DB 16; Length 2297;
Best Local Similarity 100.0%; Pred. No. 5,9e-11;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCAGTACTGCACACCGCGACATGGCGTAATCATGACGACTGTATACAGACTTGACACT 60
Db 406 GGCAGTACTGCACACCGCGACATGGCGTAATCATGACGACTGTATACAGACTTGACACT 465

RESULT 3

US-10-057-136-12
; Sequence 12, Application US/10057136
; Publication No. US20030021770A1
; GENERAL INFORMATION:
; APPLICANT: SCHLOM, JEFFREY
; APPLICANT: KANTOR, JUDITH
; APPLICANT: KUFE, DONALD
; APPLICANT: PANICALI, DENNIS
; APPLICANT: GRITZ, LINDA
; TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
; FILE REFERENCE: 700953/47113C
; CURRENT APPLICATION NUMBER: US/10/057,136
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/366,670
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: PCT/US98/03693
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/038,253
; PRIOR FILING DATE: 1997-02-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-136-12

Query Match 62.7%; Score 37.6; DB 14; Length 60;
Best Local Similarity 76.7%; Pred. No. 0.0018;

Matches 46; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 1 GGCAGTACTGCACACCGCGACATGGCGTAATCATGACGACTGTATACAGACTTGACACT 60
Db 1 GGTAGTACAGCGCACCGCACATGGCGTACAGAGCGCTCCGATACGAGACCGCGCCT 60

RESULT 4

US-10-057-136-10
; Sequence 10, Application US/10057136
; Publication No. US20030021770A1
; GENERAL INFORMATION:
; APPLICANT: SCHLOM, JEFFREY
; APPLICANT: KANTOR, JUDITH
; APPLICANT: KUFE, DONALD
; APPLICANT: PANICALI, DENNIS
; APPLICANT: GRITZ, LINDA
; TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
; FILE REFERENCE: 700953/47113C
; CURRENT APPLICATION NUMBER: US/10/057,136
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/366,670
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: PCT/US98/03693
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/038,253
; PRIOR FILING DATE: 1997-02-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-136-10

Query Match 59.0%; Score 35.4; DB 14; Length 60;
Best Local Similarity 79.2%; Pred. No. 0.01;
Matches 42; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 7 ACTGCACACCGCGACATGGCGTAATCATGACGACTGTATACAGACTTGACACT 59
Db 7 ACTGCCCTCCGCGCATGGTGTGACTGACTGCTCTGACACAGGACGACCC 59

RESULT 5

US-10-057-136-8
; Sequence 8, Application US/10057136
; Publication No. US20030021770A1
; GENERAL INFORMATION:
; APPLICANT: SCHLOM, JEFFREY
; APPLICANT: KANTOR, JUDITH
; APPLICANT: KUFE, DONALD
; APPLICANT: PANICALI, DENNIS
; APPLICANT: GRITZ, LINDA
; TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
; FILE REFERENCE: 700953/47113C
; CURRENT APPLICATION NUMBER: US/10/057,136
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/366,670
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: PCT/US98/03693
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/038,253
; PRIOR FILING DATE: 1997-02-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-136-8

Query Match 59.0%; Score 35.4; DB 14; Length 60;
Best Local Similarity 79.2%; Pred. No. 0.01;

Query Match 55.7%; Score 33.4; DB 14; Length 60;
Best Local Similarity 72.9%; Pred. No. 0.052;
Matches 43; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1 GGCAGTACTGCACACCGGACATGCGTAACTACGACCTGATCAAGACCTGCACC 59
DB 1 GGCAGACCGGACCGCCGCCGACACGGGGTCAAGGCGCGCACACTGACCTGGGCC 59

RESULT 6

US-10-057-136-9
Sequence 9, Application US/10057136
Publication No. US20030021770A1

GENERAL INFORMATION:
APPLICANT: SCHLOM, JEFFREY
APPLICANT: KANTOR, JUDITH
APPLICANT: KUFE, DONALD
APPLICANT: PANICALI, DENNIS
APPLICANT: GRITZ, LINDA
TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
FILE REFERENCE: 700953/47113C
CURRENT APPLICATION NUMBER: US/10/057,136
CURRENT FILING DATE: 2002-01-25
PRIOR APPLICATION NUMBER: 09/366,670
PRIOR FILING DATE: 1999-08-03
PRIOR APPLICATION NUMBER: PCT/US98/03693
PRIOR FILING DATE: 1998-02-24
PRIOR APPLICATION NUMBER: 60/038,253
PRIOR FILING DATE: 1997-02-24
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 9
LENGTH: 60
TYPE: DNA
ORGANISM: Homo sapiens
US-10-057-136-9

Query Match 55.7%; Score 33.4; DB 14; Length 60;
Best Local Similarity 72.9%; Pred. No. 0.052;
Matches 43; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1 GGCAGTACTGCACACCGGACATGCGTAACTACGACCTGATCAAGACCTGCACC 59
DB 1 GGAAGTACCGCTCCACCTGCACACGGGGTCAAGGCGCGCACACTGACCTGGGCC 59

RESULT 7

US-10-057-136-7
Sequence 7, Application US/10057136
Publication No. US20030021770A1

GENERAL INFORMATION:
APPLICANT: SCHLOM, JEFFREY
APPLICANT: KANTOR, JUDITH
APPLICANT: KUFE, DONALD
APPLICANT: PANICALI, DENNIS
APPLICANT: GRITZ, LINDA
TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
FILE REFERENCE: 700953/47113C
CURRENT APPLICATION NUMBER: US/10/057,136
CURRENT FILING DATE: 2002-01-25
PRIOR APPLICATION NUMBER: 09/366,670
PRIOR FILING DATE: 1999-08-03
PRIOR APPLICATION NUMBER: PCT/US98/03693
PRIOR FILING DATE: 1998-02-24
PRIOR APPLICATION NUMBER: 60/038,253
PRIOR FILING DATE: 1997-02-24
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 7
LENGTH: 60
TYPE: DNA
ORGANISM: Lactococcus lactis

TYPE: DNA
ORGANISM: Homo sapiens
US-10-057-136-7

Query Match 55.3%; Score 33.2; DB 14; Length 60;
Best Local Similarity 75.9%; Pred. No. 0.061;
Matches 41; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 7 ACTGCACACCGGACATGCGTAACTACGACCTGATCAAGACCTGCACC 60
DB 7 ACGGCCCGCTGCTCAGGTGTAACTCCGCCCGGATACGACCGGCGCT 60

RESULT 8

US-10-057-136-11
Sequence 11, Application US/10057136
Publication No. US20030021770A1

GENERAL INFORMATION:
APPLICANT: SCHLOM, JEFFREY
APPLICANT: KANTOR, JUDITH
APPLICANT: KUFE, DONALD
APPLICANT: PANICALI, DENNIS
APPLICANT: GRITZ, LINDA
TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
FILE REFERENCE: 700953/47113C
CURRENT APPLICATION NUMBER: US/10/057,136
CURRENT FILING DATE: 2002-01-25
PRIOR APPLICATION NUMBER: 09/366,670
PRIOR FILING DATE: 1999-08-03
PRIOR APPLICATION NUMBER: PCT/US98/03693
PRIOR FILING DATE: 1998-02-24
PRIOR APPLICATION NUMBER: 60/038,253
PRIOR FILING DATE: 1997-02-24
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 11
LENGTH: 60
TYPE: DNA
ORGANISM: Homo sapiens
US-10-057-136-11

Query Match 48.3%; Score 29; DB 14; Length 60;
Best Local Similarity 71.7%; Pred. No. 1.7;
Matches 38; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 7 ACTGCACACCGGACATGCGTAACTACGACCTGATCAAGACCTGCACC 59
DB 7 ACGGACCTCCGACGACGAGATCACTGTGACCCGACACCCGCTCAGCTCC 59

RESULT 9

US-10-369-493-42059/c
Sequence 42059, Application US/10369493
Publication No. US20030233675A1

GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 42059
LENGTH: 1230
TYPE: DNA
ORGANISM: Lactococcus lactis

US-10-369-493-42059

Query Match 47.7%; Score 28.6; DB 15; Length 1230;
Best Local Similarity 72.5%; Pred. No. 4.2;
Matches 37; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 10 GCACCAACCGGCGCATGGCTTAACATCAGACACCTGATCAAGACCTGCACCT 60
DB 503 GCACCAACCGGCGCATGGCTTAACATCAGACACCTGATCAAGACCTGCACCT 453

RESULT 10

US-10-437-963-15324/C
; Sequence 15324, Application US/10437963
; Publication No. US2004012343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovacic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barzak, Brad
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 15324
; LENGTH: 780
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_21178C.1
US-10-437-963-15324

Query Match 46.7%; Score 28; DB 17; Length 780;
Best Local Similarity 71.2%; Pred. No. 6.3;
Matches 37; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 8 CTGCACCAACCGGCGCATGGCTTAACATCAGACACCTGATCAAGACCTGCACCT 59
DB 360 CAGCACCAACCGGCGCATGGCTTAACATCAGACACCTGATCAAGACCTGCACCT 309

RESULT 11

US-10-057-136-6
; Sequence 6, Application US/10057136
; Publication No. US20030021770A1
; GENERAL INFORMATION:
; APPLICANT: SCHLOM, JEFFREY
; APPLICANT: KANTOR, JUDITH
; APPLICANT: KUF, DONALD
; APPLICANT: PANICALI, DENNIS
; APPLICANT: GRITZ, LINDA
; TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
; FILE REFERENCE: 700953/47113C
; CURRENT APPLICATION NUMBER: US/10/057,136
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/366,670
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: PCT/US98/03693
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/038,253
; PRIOR FILING DATE: 1997-02-24
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 6
; LENGTH: 60
; TYPE: DNA

; ORGANISM: Homo sapiens
US-10-057-136-6

Query Match 45.7%; Score 27.4; DB 14; Length 60;
Best Local Similarity 69.8%; Pred. No. 6.3;
Matches 37; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 7 ACTGACCAACCGGCGCATGGCTTAACATCAGACACCTGATCAAGACCTGCACCT 59
DB 7 ACAGCTCTCCCGCATGGGCTTAACATCAGACACCTGATCAAGACCTGCACCT 59

RESULT 12

US-10-479-638-10/C
; Sequence 10, Application US/10479638
; Publication No. US20040210956A1
; GENERAL INFORMATION:
; APPLICANT: Don A. Roth
; APPLICANT: Randolph V. Lewis
; TITLE OF INVENTION: The University of Wyoming
; FILE REFERENCE: WYO.02-0004US
; CURRENT APPLICATION NUMBER: US/10/479,638
; CURRENT FILING DATE: 2003-12-03
; PRIOR APPLICATION NUMBER: PCT/US02/18256
; PRIOR FILING DATE: 2002-06-06
; PRIOR APPLICATION NUMBER: 60/296,184
; PRIOR FILING DATE: 2001-06-06
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 1682
; TYPE: DNA
; ORGANISM: Argiope trifasciata
US-10-479-638-10

Query Match 45.7%; Score 27.4; DB 18; Length 1682;
Best Local Similarity 69.8%; Pred. No. 12;
Matches 37; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 8 CTGCACCAACCGGCGCATGGCTTAACATCAGACACCTGATCAAGACCTGCACCT 60
DB 871 CTGCTCCACCGGCGCATGGCTTAACATCAGACACCTGATCAAGACCTGCACCT 819

RESULT 13

US-10-057-136-2
; Sequence 2, Application US/10057136
; Publication No. US20030021770A1
; GENERAL INFORMATION:
; APPLICANT: SCHLOM, JEFFREY
; APPLICANT: KANTOR, JUDITH
; APPLICANT: KUF, DONALD
; APPLICANT: PANICALI, DENNIS
; APPLICANT: GRITZ, LINDA
; TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
; FILE REFERENCE: 700953/47113C
; CURRENT APPLICATION NUMBER: US/10/057,136
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/366,670
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: PCT/US98/03693
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/038,253
; PRIOR FILING DATE: 1997-02-24
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-136-2

Query Match 45.0%; Score 27; DB 14; Length 60;
Best Local Similarity 66.1%; Pred. No. 8.6;
Matches 39; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

Qy 1 GGCAGTACTGCACCGGACATGCGTAACTGACACCTGATCAAGACCTGCACC 59
|||
Db 1 GGCTCCACCGCCCCCGACCGGTGTACCTCGCCCCCGACACGAGCGGCCCC 59

RESULT 14

US-10-057-136-13

; Sequence 13, Application US/10057136
; Publication No. US20030021770A1
; GENERAL INFORMATION:

; APPLICANT: SCHLON, JEFFREY
; APPLICANT: KANTOR, JUDITH
; APPLICANT: KUFE, DONALD
; APPLICANT: PANICALI, DENNIS
; APPLICANT: GRITZ, LINDA

; TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
; FILE REFERENCE: 700953/47113C
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/366,670
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: PCT/US98/03693
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/038,253
; PRIOR FILING DATE: 1997-02-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 78
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-136-13

Query Match 45.0%; Score 27; DB 14; Length 78;
Best Local Similarity 66.1%; Pred. No. 9;
Matches 39; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

Qy 1 GGCAGTACTGCACCGGACATGCGTAACTGACACCTGATCAAGACCTGCACC 59
|||
Db 1 GGCTCCACCGACCGCCGACCGGTGTACCTCGCCCCCGACACGAGCGGCCCC 59

RESULT 15

US-10-447-839A-75/C

; Sequence 75, Application US/10447839A
; Publication No. US20040018181A1
; GENERAL INFORMATION:

; APPLICANT: KuFe, Donald W.
; APPLICANT: Kharbanda, Suresnder
; APPLICANT: Weisman, Steven D.
; TITLE OF INVENTION: MUC1 INTERFERENCE RNA COMPOSITIONS AND METHODS DERIVED THEREFROM
; FILE REFERENCE: 1000.05.009
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: 10/293,391
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 09/951,938
; PRIOR FILING DATE: 2001-09-11
; PRIOR APPLICATION NUMBER: 60/231,841
; PRIOR FILING DATE: 2000-09-11
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 75
; LENGTH: 1424
; TYPE: RNA
; ORGANISM: ARTIFICIAL
; FEATURE:

; OTHER INFORMATION: Synthesized Sequence
US-10-447-839A-75

Query Match 45.0%; Score 27; DB 16; Length 1424;
Best Local Similarity 66.1%; Pred. No. 16;
Matches 39; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

Qy 1 GGCAGTACTGCACCGGACATGCGTAACTGACACCTGATCAAGACCTGCACC 59
|||
Db 1041 GGCTCCACCGCCCCCGACCGGTGTACCTCGCCCCCGACACGAGCGGCCCC 983

Search completed: January 16, 2005, 09:30:23
Job time : 188.4 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2005, 17:40:52 : Search time 722.8 Seconds
(without alignments)
3925.542 Million cell updates/sec

Title: US-10-057-136-5

Perfect score: 60
Sequence: 1 ggatccacgcgcgcgcctgc.....ccgacacgcgcgcgcctccc 60

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 2364849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pac:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_scs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	2297	6	BD272907 A recombinant vector expressing multiple constitutively active molecules and uses thereof.
2	36.6	61.0	2135	6	AX959916 Sequence
3	35.2	58.7	1320	6	E08764 CDNA encodi
4	35	58.3	120	6	AX192396 Sequence
5	35	58.3	120	6	BD000571 Human pol
6	35	58.3	180	6	E08763 CDNA encodi
7	35	58.3	1414	12	AP423031 Synthetic
8	35	58.3	1455	6	CQ715242 Sequence
9	35	58.3	1457	6	AX959914 Sequence
10	35	58.3	1721	6	CQ711290 Sequence
11	35	58.3	1721	6	AR492306 Sequence
12	35	58.3	1721	6	AX335860 Sequence
13	35	58.3	1721	6	AX440427 Sequence
14	35	58.3	1721	6	AX587588 Sequence
15	35	58.3	1721	9	HSTBYMA Human mRNA
16	35	58.3	1774	6	AX959684 Sequence
17	35	58.3	1774	6	AX959912 Sequence
18	35	58.3	1804	6	AR492318 Sequence
19	35	58.3	1804	6	AX335367 Sequence

20	35	58.3	1804	9	HUMMUCAB	J05581 Human poly
21	35	58.3	1334	12	AF423030	AF423030 Synthetic
22	35	58.3	1835	6	AX959918	AX959918 Sequence
23	35	58.3	1971	6	AX963157	AX963157 Sequence
24	35	58.3	2037	6	AX963159	AX963159 Sequence
25	35	58.3	2238	9	HSSBTA	X52228 Human mRNA
26	35	58.3	4139	6	CO834017	CO834017 Sequence
27	35	58.3	4139	6	AX334899	AX334899 Sequence
28	35	58.3	4139	6	AX335372	AX335372 Sequence
29	35	58.3	4139	6	AX336712	AX336712 Sequence
30	35	58.3	4139	6	AX409474	AX409474 Sequence
31	35	58.3	4139	6	AX440481	AX440481 Sequence
32	35	58.3	4139	6	HUMPANMU	J05582 Human pancr
33	35	58.3	7188	9	AY463543	AY463543 Homo sapi
34	35	58.3	8181	6	AX406624	AX406624 Sequence
35	35	58.3	8181	9	HUMPEM	M61170 Human poly
36	35	58.3	8186	6	AR492320	AR492320 Sequence
37	35	58.3	133525	9	AL713999	AL713999 Human DNA
38	34.2	57.0	1258	9	AY261332	AY261332 Homo sapi
39	34.2	57.0	1935	9	AY216268	AY216268 Homo sapi
40	34.2	57.0	3420	9	HSMB04095	AL832784 Homo sapi
41	34.2	57.0	3492	9	BC042651	BC042651 Homo sapi
42	34.2	57.0	3499	9	AF445801	AF445801 Homo sapi
43	34.2	57.0	107689	9	AL354681	AL354681 Human DNA
44	33.6	56.0	491	9	HUMEPISIB1	M32739 Human epist
45	33.6	56.0	518	6	AR492402	AR492402 Sequence

ALIGNMENTS

RESULT 1	BD272907	2297 bp	DNA	linear	PAT 17-JUL-2003
LOCUS	BD272907				
DEFINITION	A recombinant vector expressing multiple constitutively active molecules and uses thereof.				
ACCESSION	BD272907	1	GI:33082675		
VERSION	BD272907				
KEYWORDS	JP 2002531133-A/1.				
SOURCE	synthetic construct				
ORGANISM	artificial sequences.				
REFERENCE	1 (bases 1 to 2297)				
AUTHORS	Schlom, J., Hodge, J. and Panticali, D.				
TITLE	A recombinant vector expressing multiple constitutively active molecules and uses thereof				
JOURNAL	Patent: JP 2002531133-A 1 24-SEP-2002;				
COMMENT	THE UNITED STATES OF AMERICA, THERION BIOLOGICS CORP				
OS	Artificial Sequence				
PN	JP 2002531133-A/1				
PD	24-SEP-2002				
PF	12-NOV-1999 JP 2000586927				
PR	09-DEC-1998 US 60/111582				
PI	JEFFREY SCHLOM, JAMES HODGE, DENNIS PANICALI				
PC	CI2N15/02, A61K35/12, A61K35/74, A61K35/76, A61K38/00, A61K39/00, A61K39/12, A61K39/21, A61K39/235, A61K39/245, A61K39/275, A61K39/29, A61K48/00.				
PC	A61P1/04, A61P31/04, A61P31/10, A61P31/12, A61P35/00, A61P37/02, PC A61P37/05.				
PC	CI2N1/16, CI2N1/19, CI2N1/21, CI2N5/10, CI2N7/00, CI2Q1/02, G01N33/53, CI2N15/00.				
PC	CI2N5/00, A61K37/02				
CC	Description of Artificial Sequence: VECTOR SEQUENCE FH				
FT	Location/Qualifiers				
FT	source				
FEATURES	Location/Qualifiers				
source	1..2297				
	/organism="Artificial Sequence".				
	/mol_type="synthetic construct"				
	/db_xref="taxon:32630"				
ORIGIN					

Query Match 100.0%; Score 60; DB 6; Length 2297;
Best Local Similarity 100.0%; Pred. No. 2,8e-05;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGATCCACCGCGCGCCCTGCGCACGAGTGAAGTGGCGCCCGACACGCGCCCGCTGCC 60
|||||
Db 466 GGATCCACCGCGCGCCCTGCGCACGAGTGAAGTGGCGCCCGACACGCGCCCGCTGCC 525

RESULT 2
LOCUS AX959916 2135 bp DNA linear PAT 14-JAN-2004
DEFINITION Sequence 21 from Patent WO03100060.
ACCESSION AX959916
VERSION AX959916.1 GI:40880144
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1
AUTHORS Burden,N.G., Ellis,J.H. and Hamblin,P.A.
TITLE Muc-1 antigen with reduced number of vnter repeat units
JOURNAL Patent: WO 03100060-A 21 04-DEC-2003;
GLAXO GROUP LIMITED (GB)

FEATURES
source location/Qualifiers
1..2135
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 61.0%; Score 36.6; DB 6; Length 2135;
Best Local Similarity 76.3%; Pred. No. 15;
Matches 45; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 1 GGATCCACCGCGCGCCCTGCGCACGAGTGAAGTGGCGCCCGACACGCGCCCGCTGCC 59
|||||
Db 642 GGCTCCACCGCGCCCGCCCGACGAGTGAAGTGGCGCCCGACACGCGCCCGCTGCC 700

RESULT 3
LOCUS E08764 1320 bp RNA linear PAT 29-SEP-1997
DEFINITION cDNA encoding C-terminal fragment of human glycoprotein 39.
ACCESSION E08764
VERSION E08764.1 GI:2176876
KEYWORDS JP 1995051065-A/2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 1320)
AUTHORS Masuzawa,Y., Muramatsu,T. and Miyuchi,T.
TITLE GLYCOPROTEIN 39 GENE
JOURNAL Patent: JP 1995051065-A 2 28-FEB-1995;
NIPPON KOTAI KENKYUSHO:KK, KAGOSHIMA UNIV
OS Homo sapiens (human)
PN JP 1995051065-A/2
PD 28-FEB-1995
PF 21-FEB-1992 JP 1992035085
PI MASUZAWA YASUSHI, MURAMATSU TAKASHI, MIYUCHI TERUO PC
C12115/09, G01N33/50//C07K14/00, C12P21/02;
CC strandedness: Double;
CC topology: linear;
CC hypothetical: No;
CC anti-sense: No;
CC Feature is identified by similarity;
FH Key Location/Qualifiers
FH source 1..1320
FT /organism="Homo sapiens"

FT /cell_type='gastric cancer'
FT /cell_line='KATO-3'
FT /clone_lib='lambda gtl1 KATO-3 CDNA library'
FT /clone='lambda KP39'
FT mat_peptide 1..981
FT /product='C-terminal fragment of human FT
FT glycoprotein 39'
FT /partial
FT polyA_signal 1267..1272
FT polyA_site 1293..1320.
FT location/Qualifiers
1..1320
/organism="Homo sapiens"
/mol_type="genomic RNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 58.7%; Score 35.2; DB 6; Length 1320;
Best Local Similarity 76.8%; Pred. No. 38;
Matches 43; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 1 GGATCCACCGCGCGCCCTGCGCACGAGTGAAGTGGCGCCCGACACGCGCCCGCTGCC 56
|||||
Db 1 GGCTCCACCGCGCCCGCCCGACGAGTGAAGTGGCGCCCGACACGCGCCCGCTGCC 56

RESULT 4
LOCUS AX192396 120 bp DNA linear PAT 15-AUG-2001
DEFINITION Sequence 2 from Patent EP1103623.
ACCESSION AX192396
VERSION AX192396.1 GI:15210363
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1
AUTHORS Taylor-Papadimitriou,T., Burchell,T. and Gendler,S.
TITLE Human mucin core protein: nucleic acid probes, peptide fragments
and antibodies thereto, and uses thereof in diagnostic and
therapeutic methods
JOURNAL Patent: EP 1103623-A 2 30-MAY-2001;
IMPERIAL CANCER RESEARCH TECHNOLOGY LIMITED (GB)

FEATURES
source location/Qualifiers
1..120
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 58.3%; Score 35; DB 6; Length 120;
Best Local Similarity 74.6%; Pred. No. 74;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 1 GGATCCACCGCGCGCCCTGCGCACGAGTGAAGTGGCGCCCGACACGCGCCCGCTGCC 59
|||||
Db 87 GGCTCCACCGCGCCCGCCCGACGAGTGAAGTGGCGCCCGACACGCGCCCGCTGCC 29

RESULT 5
LOCUS BD000571 120 bp DNA linear PAT 31-JAN-2002
DEFINITION Human polymorphic epithelial mucin core protein and nucleic acid
encoding the protein.
ACCESSION BD000571
VERSION BD000571.1 GI:18623684
KEYWORDS JP 2000333675-A/2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 120)

AUTHORS Papadimitrov,J.T., Jendora,S. and Bachieri,J.
TITLE Human polymorphic epithelial mucin core protein and nucleic acid encoding the protein
JOURNAL Patent: JP 200033675-A 2 05-DEC-2000;
 IMPERIAL CANCER RESEARCH TECHNOLOGY LTD
COMMENT OS Homo sapiens (human)
 PN JP 200033675-A/2
 PD 05-DEC-2000
 PF 26-APR-2000 JP 2000125724
 PR 07-JAN-1987 GB 8700269,07-JAN-1987 GB 8700279 PR
 22-APR-1987 US 041306,09-NOV-1987 GB 8726172 PI JOYCE
 TAYLOR PAPADIMITROV,SANDRA JENDORA,JOY BACHIERU PC
 C12N15/02,A61K38/00,A61K39/395,A61K39/00,A61P35/00, PC
 C07K14/47
 PC C07K16/44,C12N5/10,C12P21/08// (C12N15/02,C12R1:91),(C12N5/10,
 C12R1:91),
 PC C12N15/00,A61K37/02,C12N5/00,(C12N15/00,C12R1:91),(C12N5/00,
 PC C12R1:91)
 CC
 FH Key
 FT source
FEATURES
 source
 1..120
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
ORIGIN
 Query Match 58.3%; Score 35; DB 6; Length 120;
 Best Local Similarity 74.6%; Pred. No. 74;
 Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
 QY 1 GGATCCACCGCGCCGCTGCGCAGCAGTGCAGTGCAGCGCCCGCGCTCC 59
 87 GGCTCCACCGCGCCCGCCAGCCAGCGGTGCACCTCGCGCCCGAGACCGCGCGCC 29
Db
RESULT 6
LOCUS E08763 180 bp RNA linear PAT 29-SEP-1997
DEFINITION cDNA encoding N-terminal fragment of human glycoprotein 39.
ACCESSION E08763
VERSION E08763.1 GI:2176875
KEYWORDS JP 1995051065-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 180)
 Masuzawa,Y., Muramatsu,T. and Miyachi,T.
GLYCOPROTEIN 39 GENE
PATENT: JP 1995051065-A 1 28-FEB-1995;
 NIPPON KOUTAI KENKYUSHO:KK, KAGOSHIMA UNIV
OS Homo sapiens (human)
PN JP 1995051065-A/1
PD 28-FEB-1995
PF 21-FEB-1992 JP 1992035085
PI MASUZAWA YASUSHI, MURAMATSU TAKASHI, MIYACHI TERUO PC
 C12N15/09,G01N33/50//C07K14/00,C12P21/02;
 CC strandedness: Double;
 CC topology: Linear;
 CC hypothetical: No;
 CC anti-sense: No;
 CC Feature is identified by similarity;
 FH Key
 FT source
 1..180
 /organism="Homo sapiens"
 /cell_type="gastric cancer"
 /cell_line="KATO-3"
 /clone_lib="lambda gtl1 KATO-3 cDNA library"
 /clone="lambda KP39"

FT mat_peptide 1..180
FT /product="N-terminal fragment of human FT
 glycoprotein 39"
FT repeat_region 1..180
FT repeat_unit 1..60.
FT Location/Qualifiers
FEATURES
 source
 1..180
 /organism="Homo sapiens"
 /mol_type="genomic RNA"
 /db_xref="taxon:9606"
ORIGIN
 Query Match 58.3%; Score 35; DB 6; Length 180;
 Best Local Similarity 74.6%; Pred. No. 68;
 Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
 QY 1 GGATCCACCGCGCCGCTGCGCAGCAGTGCAGTGCAGCGCCCGCGCTCC 59
 121 GGCTCCACCGCGCCCGCCAGCCAGCGGTGCACCTCGCGCCCGAGACCGCGCGCC 179
Db
RESULT 7
LOCUS AF423031 1414 bp mRNA linear SYN 10-JUL-2003
DEFINITION Synthetic construct Homo sapiens mucin variant MUC1-CT58 (MUC1)
ACCESSION AF423031
VERSION AF423031.1 GI:19338621
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequence.
 1 (bases 1 to 1414)
 Hinojosa-Kurtzberg,A.M., Johansson,M.E., Madсен,C.S., Hansson,G.C.
 and Gendler,S.J.
AUTHORS Novel MUC1 splice variants contribute to mucin overexpression in
 CPTN-deficient mice
TITLE Am. J. Physiol. Gastrointest. Liver Physiol. 284 (5), G853-G862
 (2003)
JOURNAL MEDLINE 22570517
PUBMED 12529261
REFERENCE 2 (bases 1 to 1414)
AUTHORS Hinojosa-Kurtzberg,A.M. and Gendler,S.J.
TITLE Direct Submission
JOURNAL Submitted (21-SEP-2001) Biochemistry, Mayo Clinic, 13400 E. Shea
 Boulevard, Scottsdale, AZ 85259, USA
FEATURES
 source
 1..1414
 /organism="Synthetic construct"
 /mol_type="mRNA"
 /db_xref="taxon:32630"
 /note="Homo sapiens gene in transgenic Mus musculus
 C57BL/6; isolated from intestinal mucosa"
 1..1414
 /gene="MUC1"
 1..1386
 /gene="MUC1"
 /note="alternatively spliced; contains exon 6b resulting
 in variant carboxy-terminal domain; lacks sites for
 beta-catenin and Grb2 interactions; derived from Homo
 sapiens"
 /codon_start=1
 /transl_table=11
 /product="mucin variant MUC1-CT58"
 /protein_id="AA186735.1"
 /db_xref="GI:19338622"
 /translation="MTPTGSPFLLLLTLVTYTGSGHASTPGEKESATQSS
 VPSSTKNAVSMSTSVLSHSPGSSSTTQODVTLATPESASATWGDVTV
 VTPALGISTTPPAHDVTSAPDNKAPGSAIPAHGVTSAPDTPRPAAGSAIPAHGVTS
 AADNPALGISTAPVPAHNTSASGASGASTLVHNGTSARATTPPASKSTPSPSH
 SDPTPLASHSTKTSASTHSTVPLTSSNHSSTPQSTGVSPPFLSFHLSNLFNS
 SLEDSSTIYQELQDISEMFLQIKKGSGFLGSLNKKRPSGVVQTLAPREGTIV

ORIGIN

Query Match	58.3%	Score 35	DB 12	Length 1414
Best Local Similarity	74.6%	Pred. No. 41		
Matches 44	Conservative	0	Mismatches 15	Indels 0
				Gaps 0

Qy 1 GGATCCACCGCGCCCGCTGCGCAGGAGTGAAGTGGCGCCGAGACGCGCCCGCTCC 59
|||
Db 385 GGCTCACCGGCCCGCCGAGCCCAAGGTGACCTCGGCCCGAGACCAAGCCGCGCCCC 443

RESULT	8
CQ715242	
LOCUS	CQ715242
DEFINITION	Sequence 1176 from Patent WO20068579.
ACCESSION	CQ715242
VERSION	CQ715242.1 GI:42276099
KEYWORDS	.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens

REFERENCE	AUTHORS	TITLE
1	Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.	kits, such as nucleic acid arrays, comprising a majority of

JOURNAL
Patent: WO 02068579-A 1176 06-SEP-2002;
PE Corporation (NY) (US)

```
source
1. .1455
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
```

Query Match	58.3%	Score 35;	DB 6;	Length 1455;
Best Local Similarity	74.6%	Pred. No. 41;		
Matches 44;	Conservative	0;	Mismatches 15;	Indels 0;
			Gaps	0

Dy
1 GGATCCACCGCGGCGCTGGCGCAGGAGTGAAGTCGCGGCCCAACGGGCCCCGTCC 59

Dd
412 GGTCCAACGCCCCCCCAGGCCAAGTGTCACCTCGGCCCCGACACAGAAGCCGCCCC 470

LOCUS	SEQUENCE	FROM PATENT	MO3100060.	DNA	1457 bp	LINEAR	PAT 14-JAN-2006
AX959914	AX959914	19	GI:40860143				

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1

GLAXO GROUP LIMITED (GB)
Patent: WO 0310060-A 19 04-DEC-2003;
Muc-1 antigen with reduced number of vnt_r repeat units
TITLE
JOURNAL
Elliott, N.G.; Ellis, D.H. and Hamblin, P.A.
Cancer, N.G.; Ellis, D.H. and Hamblin, P.A.

```

FEATURES
SOURCE      location/Qualifiers
            1..1457
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"
ORIGIN

```

Query Match	58.3%;	Score 35;	DB 6;	Length 1457;
Best Local Similarity	74.6%;	Pred. NO. 41;		

Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 1 GGATCCACCGCGGCGCTGGGCAGAGTGAAGTCGGCGGCCGACAGCGCCCGCTCC 59
 |||||
Db 403 GGCTCCACCGCCCCCAGCCCAAGGTGTCACTCGCCCGGACACCAAGGCGGCCCC 461

RESULT 10				
CQ771290				
LOCUS	CQ771290	1721 bp	DNA	linear
DEFINITION	Sequence 1 from Patent WO2004009632.			PAT 04-MAR-2004
ACCESSION	CQ771290			
VERSION	CQ771290.1	GI:45125397		
KEYWORDS				

ORGANISM	REFERENCE
Homo sapiens	1
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	

TITLE Method for the production of an immunostimulating mucin (muc1)
JOURNAL Patent: WO 2004009632-A 1 29-JAN-2004;
Nemod Immuntherapie AG (DE)

```
/organism="Homo sapiens"  
/mol_type="unassigned DNA"  
/db_xref="taxon:9606"
```

ORIGIN

Query Match	58.3%	Score 35;	DB 6;	Length 1721;
Best Local Similarity	74.6%;	Pred. No. 39;		
Matches 44;	Conservative	0;	Mismatches 15;	Indels 0;
			Gaps	0

DY 1 GGATCCACCGGCGCCCTGTGGCAGGAGTGA CTTGGGGCCCCGACACGGCCCCCGCTCC 59

Db 562 GGCTTCACCGCGCCCGAGCCCA CGGTGCACCTCGGCCCGGACACCAGCCGCCCC 620

RESULT 11			
AR492306	AR492306	1721 bp	DNA
LOCUS	Sequence 3	from patent US 6716667.	linear
DEFINITION	AR492306		
ACCESSION	AR492306.1	GI:47260880	
VERSION			
FEATURES			

SOURCE	ORGANISM	Unknown.
	Unknown.	

REFERENCE 1 (bases 1 to 1721)

TITLE	Antisense modulation of mucin 1, transmembrane expression
JOURNAL	Patent: US 6716627-A 3 06-APR-2004;
REFERENCE	

```
source      1..1721
            /organism="unknown"
            /mol_type="genomic DNA"
            DBIN
```

ORIGIN

Query Match	58.3%	Score 35;	DB 6;	Length 1721;
Best Local Similarity	74.6%;	Pred. No. 39;		
Matches 44;	Conservative	0;	Mismatches 15;	Indels 0;
				Gaps 0;

QY 1 GGATCCACCGCGCGCTGCGCACGAGTACGTGGCGCGCCGACACGCGCCCGCTCC 59

Db 562 GGCCTCACCGCGCGCGCACGAGTGTACACTCGGCGCCGGAACACGAGCGCGCCC 620

RESULT 12			
AX335860			
LOCUS	AX335860	1721 bp	DNA
DEFINITION	Sequence 6369 from Parent W00194629.		linear
ACCESSION	AX335860		PAT 09-JAN-2002

VERSION AX35860.1 GI:18126579
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS 1 Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G., Horrigan, S., Soppet, D.R. and Weaver, Z.
TITLE Cancer gene determination and therapeutic screening using signature gene sets
JOURNAL Patent: WO 0194629-A 6369 13-DEC-2001;
Avalon Pharmaceuticals (US)
FEATURES Location/Qualifiers
source 1..1721
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 58.3%; Score 35; DB 6; Length 1721;
Best Local Similarity 74.6%; Pred. No. 39;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 1 GGATCCACCGCGCGCTGCGACGAGTGACGTGGCGCCGACACGCGCCCGCTCC 59
|||
562 GGCTCCACCGCGCGCGCTGCGACGAGTGACGTGGCGCCGACACGCGCCCGCTCC 620
Db
RESULT 13
AX440427 1721 bp DNA linear PAT 28-JUN-2002
LOCUS
DEFINITION Sequence 280 from Patent WO0190154.
ACCESSION AX440427
VERSION AX440427.1 GI:21665237
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS 1 Xu, J., Mitcham, J.L., Harlocker, S.L., Dillon, D.C., Secrist, H., Lodes, M.J., Algate, P.A., Fling, S.P., Mannion, J., Benson, D.R. and Carter, D.
TITLE Compositions and methods for the therapy and diagnosis of ovarian cancer
JOURNAL Patent: WO 0190154-A 280 29-NOV-2001;
CORIXA CORPORATION (US)
FEATURES Location/Qualifiers
source 1..1721
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 58.3%; Score 35; DB 6; Length 1721;
Best Local Similarity 74.6%; Pred. No. 39;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 1 GGATCCACCGCGCGCTGCGACGAGTGACGTGGCGCCGACACGCGCCCGCTCC 59
|||
562 GGCTCCACCGCGCGCGCTGCGACGAGTGACGTGGCGCCGACACGCGCCCGCTCC 620
Db
RESULT 14
AX587588 1721 bp DNA linear PAT 10-JAN-2003
LOCUS
DEFINITION Sequence 58 from Patent WO0246467.
ACCESSION AX587588
VERSION AX587588.1 GI:28212336
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct

artificial sequences.
REFERENCE 1 Bertucci, F., Houlgate, R., Birnbaum, D., Nguyen, C., Viens, P. and Fert, V.
AUTHORS Gene expression profiling of primary breast carcinomas using arrays of candidate genes
JOURNAL Patent: WO 0246467-A 58 13-JUN-2002;
Ipsogen (FR)
FEATURES Location/Qualifiers
source 1..1721
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
misc_feature 1..1721
/note="mucin 1, transmembrane (MUC1) gene."
ORIGIN
Query Match 58.3%; Score 35; DB 6; Length 1721;
Best Local Similarity 74.6%; Pred. No. 39;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 1 GGATCCACCGCGCGCTGCGACGAGTGACGTGGCGCCGACACGCGCCCGCTCC 59
|||
562 GGCTCCACCGCGCGCGCTGCGACGAGTGACGTGGCGCCGACACGCGCCCGCTCC 620
Db
RESULT 15
HSTXYMA 1721 bp mRNA linear PRI 14-AUG-1995
LOCUS
DEFINITION Human mRNA for transmembrane epithelial tumour mucin antigen.
ACCESSION X52229
VERSION X52229.1 GI:37053
KEYWORDS Transmembrane protein; tumor antigen.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS 1 Harsanyi, M., Tsarfaly, I., Zaretsky, J., Kotkes, P., Horev, J., Zrihan, S., Weiss, M., Green, S., Lathe, R., Keydar, I. and Wreschner, D.H.
TITLE A transcribed gene, containing a variable number of tandem repeats, codes for a human epithelial tumor antigen. cDNA cloning, expression of the transcribed gene and over-expression in breast cancer tissue
JOURNAL Eur. J. Biochem. 189 (3), 475-486 (1990)
MEDLINE 90276414
PUBMED 2112460
REFERENCE 2 (bases 1 to 1721)
AUTHORS Wreschner, D.H., Harsanyi, M., Tsarfaly, I., Smorodinsky, N., Horev, J., Zaretsky, J., Kotkes, P., Weiss, M., Lathe, R., Dion, A.S. and Keydar, I.
TITLE Human epithelial tumor antigen cDNA sequences. Differential splicing may generate multiple protein forms
JOURNAL Eur. J. Biochem. 189 (3), 463-473 (1990)
MEDLINE 90276413
PUBMED 2351132
REFERENCE 3 (bases 1 to 1721)
AUTHORS Wreschner, D.H.
TITLE Direct Submision
JOURNAL Submitted (16-MAR-1990) Wreschner D.H., Tel Aviv University, Dept of Microbiology, Faculty of Life Sciences, Tel Aviv 69978, Israel
COMMENT See <X52228> for secreted tumour antigen.
FEATURES Location/Qualifiers
source 1..1721
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="1q21-24"
/cell_line="T47D breast tumour"
/clone_11b="lambda gt11"
58..1605
CDS

Query Match 45.0%; Score 27; DB 4; Length 1804;
Best Local Similarity 66.1%; Pred. No. 1.8;
Matches 39; Conservative 0; Mismatches 20; Indels 0; Gaps 0;
QY 1 GGCAGTACTGCACACCGGCACATGGCGTAAATCAGCAGCTGTATACAAAGACTTCGACC 59
DB 457 GGCTCCACCGCCCCCGCCACCGACGCTGTACCTGGCCCCCGACACAGGCGGCCCC 515

RESULT 3
US-10-029-517-19
; Sequence 19, Application US/10029517
; Patent No. 6716227
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Susan J. Myers
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 19
; LENGTH: 8186
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 6899
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 7155
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 7184
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 7957
; OTHER INFORMATION: unknown
; NAME/KEY: intron
; LOCATION: (2997)...(3498)
; OTHER INFORMATION: intron 1
; NAME/KEY: intron:exon junction
; LOCATION: (3498)...(3499)
; OTHER INFORMATION: intron 1:exon 2
; NAME/KEY: exon
; LOCATION: (3508)...(3599)
; OTHER INFORMATION: exon 2d
; NAME/KEY: exon:intron junction
; LOCATION: (3982)...(3983)
; OTHER INFORMATION: exon 2a:intron 2a
; NAME/KEY: intron:exon junction
; LOCATION: (4205)...(4206)
; OTHER INFORMATION: intron 2c:exon 3c
; NAME/KEY: intron:exon junction
; LOCATION: (4259)...(4260)
; OTHER INFORMATION: intron 2d:exon 3d
; NAME/KEY: exon
; LOCATION: (4360)...(4328)
; OTHER INFORMATION: exon 3d
; NAME/KEY: intron:exon junction
; LOCATION: (4632)...(4633)
; OTHER INFORMATION: intron 3:exon 4
; NAME/KEY: exon
; LOCATION: (4914)...(5035)
; OTHER INFORMATION: exon 5
; NAME/KEY: intron
; LOCATION: (5266)...(6293)
; OTHER INFORMATION: intron 6
US-10-029-517-19

Query Match 45.0%; Score 27; DB 4; Length 8186;
Best Local Similarity 66.1%; Pred. No. 2.6;
Matches 39; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 1 GGCAGTACTGCACACCGGCACATGGCGTAAATCAGCAGCTGTATACAAAGACTTCGACC 59
DB 3825 GGCTCCACCGCCCCCGCCACCGACGCTGTACCTGGCCCCCGACACAGGCGGCCCC 3883

RESULT 4
US-09-016-434-1129/C
; Sequence 1129, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HERewith
; CLASSIFICATION:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1129:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3722 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g1339211
US-09-016-434-1129

Query Match 44.0%; Score 26.4; DB 4; Length 3722;
Best Local Similarity 65.0%; Pred. No. 3.5;
Matches 39; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 1 GGCAGTACTGCACACCGGCACATGGCGTAAATCAGCAGCTGTATACAAAGACTTCGACC 60
DB 2384 GGTCGCGACGCCCCCTCCGACCTTGTGGCCACACACACCTGGGAGAGCCCTTCACTT 2325

RESULT 5
US-08-464-700-1/C
; Sequence 1, Application US/08464700
; Patent No. 6232458
; GENERAL INFORMATION:
; APPLICANT: WEISS, ANTHONY S
; APPLICANT: MARTIN, STEPHEN L
; TITLE OF INVENTION: SYNTHETIC POLYNUCLEOTIDES
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:

```
ADDRESSER: Howson and Howson
STREET: Spring House Corporate Cntr, PO Box 457
CITY: Spring House
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19477

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/464,700
FILING DATE: 7-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PL6520
FILING DATE: 22-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PL9661
FILING DATE: 28-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU93/00655
FILING DATE: 16-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: GRCJUSA
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-9200
TELEFAX: 215-540-5818
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2210 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
HYPOTHETICAL: YES
ANTI-SENSE: NO
US-08-464-700-1

Query Match
Best Local Similarity 43.3%; Score 26; DB 3; Length 2210;
Matches 35; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 10 GCACCAACCGGACATGCGCTAATGACGACCTGTATACAAAGCTGCACC 59
Db 1459 GCACACCTACACCTGCTGCAACGCCAACACCCGGTGCCAGACCTACGCC 1410

RESULT 6
US-09-270-767-10939
Sequence 10939, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 10939
LENGTH: 489
TYPE: DNA
ORGANISM: Drosophila melanogaster
FEATURE:
OTHER INFORMATION: n means any nucleotide
US-09-270-767-10939

Query Match
Best Local Similarity 43.0%; Score 25.8; DB 4; Length 489;
Matches 36; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
```

```
Qy 8 CTCACACCGGACGATGGCGTAACTGACGACCTGTATACAGACCTGCACCT 60
Db 304 CAGCAACATCAGACGCTGACGACCAATCAGCAGCTGACAGCAATCAGACGCT 356

RESULT 7
US-10-029-517-101
Sequence 101, Application US/10029517
Patent No. 6716627
GENERAL INFORMATION:
APPLICANT: Kenneth W. Doble
TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
FILE REFERENCE: RTS-0352
CURRENT APPLICATION NUMBER: US/10/029,517
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 107
SEQ ID NO 101
LENGTH: 518
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION:
US-10-029-517-101

Query Match
Best Local Similarity 42.7%; Score 25.6; DB 4; Length 518;
Matches 37; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 1 GGCAGTACTGACACCGGACATGCGCTAATGACGACCTGTATACAAAGCTGC 56
Db 462 GGCTCCACCGCCGCCCGGACGCGGTGTACCTGGCCCGGACACCGGCGCGC 517

RESULT 8
US-10-029-517-16
Sequence 16, Application US/10029517
Patent No. 6716627
GENERAL INFORMATION:
APPLICANT: Kenneth W. Doble
TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
FILE REFERENCE: RTS-0352
CURRENT APPLICATION NUMBER: US/10/029,517
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 107
SEQ ID NO 16
LENGTH: 981
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: exon:exon junction
LOCATION: (464)...(465)
OTHER INFORMATION: exon 3b:exon 4
US-10-029-517-16

Query Match
Best Local Similarity 42.7%; Score 25.6; DB 4; Length 981;
Matches 37; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 1 GGCAGTACTGACACCGGACATGCGCTAATGACGACCTGTATACAAAGCTGC 56
Db 21 GGCTCCACCGCCGCCCGGACGCGGTGTACCTGGCCCGGACACCGGCGCGC 76

RESULT 9
US-10-029-517-3
Sequence 3, Application US/10029517
Patent No. 6716627
GENERAL INFORMATION:
APPLICANT: Kenneth W. Doble
TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
```

FILE REFERENCE: RTS-0352
CURRENT APPLICATION NUMBER: US/10/029,517
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 107
SEQ ID NO 3
LENGTH: 1721
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (58) ... (1605)
US-10-029-517-3

Query Match 42.7%; Score 25.6; DB 4; Length 1721;
Best Local Similarity 66.1%; Pred. No. 5.5;
Matches 37; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 1 GGCAGTACTGCACCGGACATGGCGTAACATGACCACTGATACAGACCTGC 56
DB 622 GGCTCCACCGCGCCCGCCAGCCCATGTGTCACTCGCGCCGACACAGGCCCCG 677

RESULT 10
US-10-029-517-102
Sequence 102, Application US/10029517
Patent No. 6716627
GENERAL INFORMATION:
APPLICANT: Kenneth W. Dobie
TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
FILE REFERENCE: RTS-0352
CURRENT APPLICATION NUMBER: US/10/029,517
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 107
SEQ ID NO 102
LENGTH: 3343
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
US-10-029-517-102

Query Match 42.7%; Score 25.6; DB 4; Length 3343;
Best Local Similarity 66.1%; Pred. No. 6.6;
Matches 37; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 1 GGCAGTACTGCACCGGACATGGCGTAACATGACCACTGATACAGACCTGC 56
DB 1728 GGCTCCACCGCGCCCGCCAGCCCATGTGTCACTCGCGCCGACACAGGCCCCG 1783

RESULT 11
US-10-029-517-18
Sequence 18, Application US/10029517
Patent No. 6716627
GENERAL INFORMATION:
APPLICANT: Kenneth W. Dobie
TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
FILE REFERENCE: RTS-0352
CURRENT APPLICATION NUMBER: US/10/029,517
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 107
SEQ ID NO 18
LENGTH: 572
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (67) ... (572)
US-10-029-517-18

Query Match 42.3%; Score 25.4; DB 4; Length 572;
Best Local Similarity 64.4%; Pred. No. 4.9;

Matches 38; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 1 GGCAGTACTGCACCGGACATGGCGTAACATGACCACTGATACAGACCTGC 59
DB 478 GGCTCCACCGCGCCCGCCAGCCCATGTGTCACTCGCGCCGACACAGGCCCCG 536

RESULT 12
US-09-328-352-2989/c
Sequence 2989, Application US/09328352
Patent No. 6562958
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
FILE REFERENCE: GTC99-03PA
CURRENT APPLICATION NUMBER: US/09/328,352
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 2989
LENGTH: 1260
TYPE: DNA
ORGANISM: Acinetobacter baumannii
US-09-328-352-2989

Query Match 41.7%; Score 25; DB 4; Length 1260;
Best Local Similarity 64.9%; Pred. No. 8.4;
Matches 37; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 4 AGTACTGCACCGGACATGGCGTAACATGACCACTGATACAGACCTGCACCT 60
DB 470 AGAATTTCACCCCATGACCTGTGTCACTGACACGACCGGTTAAAGTAATGAACCT 414

RESULT 13
US-09-175-928-41
Sequence 41, Application US/09175928A
Patent No. 6312821
GENERAL INFORMATION:
APPLICANT: Jacobs, Kenneth
APPLICANT: McCoy, John M.
APPLICANT: LaVallie, Edward R.
APPLICANT: Collins-Racie, Lisa A.
APPLICANT: Evans, Cheryl
APPLICANT: Metberg, David
APPLICANT: Treacy, Maurice
APPLICANT: Mi, Sha
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
FILE REFERENCE: 6006B.AJ172A
CURRENT APPLICATION NUMBER: US/09/175,928A
CURRENT FILING DATE: 1998-10-20
NUMBER OF SEQ ID NOS: 62
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 41
LENGTH: 1351
TYPE: DNA
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: unsure
LOCATION: (134)
US-09-175-928-41

Query Match 41.3%; Score 24.8; DB 3; Length 1351;
Best Local Similarity 67.3%; Pred. No. 10;
Matches 35; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 7 ACTGCACACCGGACATGGCGTAACATGACCACTGATACAGACCTGCAC 58
DB 825 ACTGTATCCCGGACATGACGACGACGACCTGTGTGTAAGATTTTCAC 876

RESULT 14


```
US-09-489-039A-3117/c
; Sequence 3117, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709, 2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 3117
; LENGTH: 1362
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-3117

Query Match          41.3%; Score 24.8; DB 4; Length 1362;
Best Local Similarity 67.3%; Pred. No. 10;
Matches 35; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY      8 CTGCACCAACGGCGCATGCGTAACATCAGCACCTGTATCAAGACCTGCACC 59
DB      1215 CCGCTGCCGACCACTAGTACACACGCGCGGATTCACGCCCTGCATC 1164

RESULT 15
US-09-765-400-64/c
; Sequence 64, Application US/09765400
; Patent No. 6691568
; GENERAL INFORMATION:
; APPLICANT: Ghazal, Peter
; APPLICANT: Huang, Huang
; TITLE OF INVENTION: Generation of Human Cytomegalovirus Yeast Artificial Chromosome
; TITLE OF INVENTION: Recombinants
; FILE REFERENCE: 98,299
; CURRENT APPLICATION NUMBER: US/09/765,400
; CURRENT FILING DATE: 2000-11-03
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 64
; LENGTH: 229354
; TYPE: DNA
; ORGANISM: Human cytomegalovirus;
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Human cytomegalovirus strain AD169 (GenBank X17403.1)
US-09-765-400-64

Query Match          41.3%; Score 24.8; DB 4; Length 229354;
Best Local Similarity 67.3%; Pred. No. 39;
Matches 35; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY      8 CTGCACCAACGGCGCATGCGTAACATCAGCACCTGTATCAAGACCTGCACC 59
DB      132654 CTGAGCCAGACCCCATCGATCTTCTGAGCAGATTCAGAACTGCACC 132603

Search completed: January 16, 2005, 03:01:16
Job time : 37.2 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2005, 19:43:55 : Search time 1550.6 Seconds
(without alignments)
1410.024 Million cell updates/sec

Title: US-10-057-136-4

Perfect score: 60
Sequence: 1 GGCAGTACGACACACACGCGC.....CTGATACAGACTGACACT 60

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_g881:*
9: gb_g882:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	29.2	48.7	107	7	CK484802 rswa0_00
C 2	29.2	48.7	148	7	CK530426 rswg0_00
C 3	29.2	48.7	247	7	CK543078 rswb0_00
C 4	29.2	48.7	350	5	BP127568 BP127568
C 5	29.2	48.7	357	7	CK541226 rswb0_00
C 6	29.2	48.7	366	7	CK542960 rswb0_00
C 7	29.2	48.7	370	5	BP127555 BP127555
C 8	29.2	48.7	375	5	CK531267 rswg0_00
C 9	29.2	48.7	375	7	CK542714 rswb0_00
C 10	29.2	48.7	378	7	CK487724 rswb0_00
C 11	29.2	48.7	379	7	CK487032 rswb0_00
C 12	29.2	48.7	393	7	CK530060 rswg0_00
C 13	29.2	48.7	400	7	CK541451 rswb0_00
C 14	29.2	48.7	407	7	CK501702 rswc0_00
C 15	29.2	48.7	417	7	CK488576 rswb0_00
C 16	29.2	48.7	434	7	CK530111 rswg0_00
C 17	29.2	48.7	443	1	AU183191 AU183191
C 18	29.2	48.7	445	7	CK531594 rswg0_00
C 19	29.2	48.7	447	7	CK532680 rswg0_01
C 20	29.2	48.7	452	7	CK530081 rswg0_00
C 21	29.2	48.7	475	7	CK542102 rswb0_00
C 22	29.2	48.7	487	7	CK529960 rswg0_00
C 23	29.2	48.7	491	7	CK538321 rswb0_00
C 24	29.2	48.7	500	5	BP127928 BP127928

C 25	29.2	48.7	500	7	CK491392 rswb0_00
C 26	29.2	48.7	510	5	BP127913 BP127913
C 27	29.2	48.7	516	7	CK485119 rswb0_00
C 28	29.2	48.7	525	7	CK541616 rswb0_00
C 29	29.2	48.7	527	7	CK530596 rswg0_00
C 30	29.2	48.7	535	7	CK540582 rswb0_00
C 31	29.2	48.7	537	7	CK532628 rswg0_01
C 32	29.2	48.7	540	7	CK538162 rswg0_00
C 33	29.2	48.7	547	7	CK486345 rswb0_00
C 34	29.2	48.7	549	7	CK491625 rswb0_01
C 35	29.2	48.7	552	7	CK540972 rswb0_00
C 36	29.2	48.7	570	7	CK542381 rswb0_00
C 37	29.2	48.7	571	7	CK545577 rswb0_01
C 38	29.2	48.7	581	7	CK532708 rswg0_01
C 39	29.2	48.7	587	5	BP127701 BP127701
C 40	29.2	48.7	589	7	CK485857 rswb0_00
C 41	29.2	48.7	604	7	CK490090 rswb0_00
C 42	29.2	48.7	604	7	CK490598 rswb0_00
C 43	28.8	48.0	568	5	BX611392 BX611392
C 44	28.6	47.7	502	7	D16065 D16065
C 45	28.6	47.7	528	1	AU090559 AU090559

ALIGNMENTS

RESULT 1
LOCUS CK484802 107 bp mRNA linear EST 14-JAN-2004
DEFINITION rswa0_001632.y1 swa Bombyx mori cDNA, mRNA sequence.
ACCESSION CK484802
VERSION CK484802.1 GI:40840561

KEYWORDS EST.
SOURCE Bombyx mori (domestic silkworm)

ORGANISM Bombyx mori

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Diptera; Bombycoidea; Bombycidae; Bombyx.

REFERENCE 1 (bases 1 to 107)

AUTHORS Xia, Q., Zhou, Z., Lu, C., Cheng, D., Dai, F., Li, B., Zhao, P., Zhu, X., Cheng, T., Chai, C., Pan, G., Xu, J., Liu, C., Lin, Y., Qian, J., Hou, Y., Wu, Z., Li, G., Pan, M., Li, C., Shen, Y., Lan, X., Yuan, L., Yang, G., Man, Y., Zhu, Y., Yu, M., Shen, W., Wu, D., Xiang, Z., Yu, J., Wang, J., Li, R., Li, H., Li, G., Su, J., Wang, X., Li, G., Zhang, Z., Li, J., Zhang, Q., Wei, N., Xu, J., Sun, H., Dong, L., Liu, D., Zhao, S., Zhao, X., Meng, Q., Lan, F., Fang, L., Li, Y., Wang, J., Ye, J., Wong, G. K.-S., Yang, H., Shi, J., Wu, Q., Li, C., Li, D., Sun, Y., Zhang, Z., Yang, Z., Huang, Y., Xi, Y., Qi, Q., He, D., Huang, H., Zhang, X., Wang, Z., Li, W. and Cao, Y.

TITLE A draft sequence for the genome of the domesticated silkworm (Bombyx mori)
JOURNAL Unpublished (2004)
COMMENT Contact: Yang Huan Ming
Beijing Genomics Institute
Chinese Academy of Sciences
Beijing Airport Industrial Zone B6, Beijing 101300, P.R.China
Email: hyan@igbp.ac.cn

FEATURES

source 1..107
/organism="Bombyx mori"
/mol_type="mRNA"
/strain="Dazhao (P50)"
/db_xref="taxon:7091"
/sex="mixed"
/tissue_type="silk gland"
/dev_stage="5th-instar day-3 larva"
/clone_lib="swa"
/note="Vector: pBluescript II SK(+)"

ORIGIN

Query Match 48.7%; Score 29.2; DB 7; Length 107;
Best local Similarity 74.0%; Pred. No. 19;
Matches 37; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 11 CACCACCGGCACATGCGCTAACTAGACCTGATACAGACCTGCACCT 60
DB 103 CAGCACCGGCACCTGAACACGACACGACCTGAACACGACACCGGCACCT 54

RESULT 2
CKS30426/c
LOCUS CKS30426
DEFINITION rswga0.002879.y1 swg Bombyx mori cDNA, mRNA sequence.
ACCESSION CKS30426
VERSION CKS30426.1 GI:40914879
KEYWORDS EST.
SOURCE Bombyx mori (domestic silkworm)
ORGANISM Bombyx mori
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
Bombycoidea; Bombycidae; Bombyx.

REFERENCE
AUTHORS Xia,Q., Zhou,Z., Lu,C., Cheng,D., Dai,F., Li,B., Zhao,P., Zha,X.,
Cheng,T., Chai,C., Pan,G., Xu,J., Liu,C., Lin,Y., Qian,J., Hou,Y.,
Wu,Z., Li,G., Pan,M., Li,C., Shen,Y., Lan,X., Yuan,L., Yang,G.,
Wan,Y., Zhu,Y., Yu,M., Shen,W., Wu,D., Xiang,Z., Yu,J., Wang,J.,
Li,R., Li,H., Li,G., Su,J., Wang,X., Li,G., Zhang,Z., Li,J.,
Zhang,Q., Wei,N., Xu,J., Sun,H., Dong,L., Liu,D., Zhao,S., Zhao,X.,
Yang,Q., Lan,F., Pang,L., Li,Y., Wang,J., Ye,J., Wong,G.K.-S.,
Mang,O., Shi,J., Wu,Q., Li,C., Li,D., Sun,Y., Zhang,Z., Yang,Z.,
Huang,Y., Xi,Y., Qi,Q., He,D., Huang,H., Zhang,X., Wang,Z., Li,W.
and Cao,Y.
A draft sequence for the genome of the domesticated silkworm
(Bombyx mori)
Unpublished (2004)
Contact: Yang Huan Ming
Beijing Genomics Institute
Chinese Academy of Sciences
Beijing Airport Industrial Zone B6, Beijing 101300, P.R.China
Email: hyang@igtp.ac.cn.
Location/Qualifiers
1.148
/organism="Bombyx mori"
/mol_type="mRNA"
/strain="Dazhao(P50)"
/db_xref="taxon:7091"
/sex="male"
/tissue_type="testis"
/dev_stage="5th-instar day-3 larva"
/clone_lib="swg"
/note="Vector: pBluescript II SK(+)"

ORIGIN
Query Match 48.7%; Score 29.2; DB 7; Length 148;
Best Local Similarity 74.0%; Pred. No. 20;
Matches 37; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 11 CACCACCGGCACATGCGCTAACTAGACCTGATACAGACCTGCACCT 60
DB 138 CAGCACCGGCACCTGAACACGACACGACCTGAACACGACACCGGCACCT 89

RESULT 3
CKS43078
LOCUS CKS43078
DEFINITION rswb0.009657.y1 swb Bombyx mori cDNA, mRNA sequence.
ACCESSION CKS43078
VERSION CKS43078.1 GI:40927532
KEYWORDS EST.
SOURCE Bombyx mori (domestic silkworm)
ORGANISM Bombyx mori
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
Bombycoidea; Bombycidae; Bombyx.

REFERENCE
AUTHORS Xia,Q., Zhou,Z., Lu,C., Cheng,D., Dai,F., Li,B., Zhao,P., Zha,X.,

Cheng,T., Chai,C., Pan,G., Xu,J., Liu,C., Lin,Y., Qian,J., Hou,Y.,
Wu,Z., Li,G., Pan,M., Li,C., Shen,Y., Lan,X., Yuan,L., Yang,G.,
Wan,Y., Zhu,Y., Yu,M., Shen,W., Wu,D., Xiang,Z., Yu,J., Wang,J.,
Li,R., Li,H., Li,G., Su,J., Wang,X., Li,G., Zhang,Z., Li,J.,
Zhang,Q., Wei,N., Xu,J., Sun,H., Dong,L., Liu,D., Zhao,S., Zhao,X.,
Yang,Q., Lan,F., Pang,L., Li,Y., Wang,J., Ye,J., Wong,G.K.-S.,
Mang,O., Shi,J., Wu,Q., Li,C., Li,D., Sun,Y., Zhang,Z., Yang,Z.,
Huang,Y., Xi,Y., Qi,Q., He,D., Huang,H., Zhang,X., Wang,Z., Li,W.
and Cao,Y.
A draft sequence for the genome of the domesticated silkworm
(Bombyx mori)
Unpublished (2004)
Contact: Yang Huan Ming
Beijing Genomics Institute
Chinese Academy of Sciences
Beijing Airport Industrial Zone B6, Beijing 101300, P.R.China
Email: hyang@igtp.ac.cn.
Location/Qualifiers
1.247
/organism="Bombyx mori"
/mol_type="mRNA"
/strain="Dazhao(P50)"
/db_xref="taxon:7091"
/sex="female"
/tissue_type="ovary"
/dev_stage="5th-instar day-3 larva"
/clone_lib="swb"
/note="Vector: pBluescript II SK(+)"

ORIGIN
Query Match 48.7%; Score 29.2; DB 7; Length 247;
Best Local Similarity 74.0%; Pred. No. 21;
Matches 37; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 11 CACCACCGGCACATGCGCTAACTAGACCTGATACAGACCTGCACCT 60
DB 8 CAGCACCGGCACCTGAACACGACACGACCTGAACACGACACCGGCACCT 57

RESULT 4
BP127568
LOCUS BP127568
DEFINITION BP127568 tees Bombyx mori cDNA clone tesso720, mRNA sequence.
ACCESSION BP127568
VERSION BP127568.1 GI:29560621
KEYWORDS EST.
SOURCE Bombyx mori (domestic silkworm)
ORGANISM Bombyx mori
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
Bombycoidea; Bombycidae; Bombyx.

REFERENCE
AUTHORS Mita,K., Moritomo,M., Shimada,T., Okano,K. and Maeda,S.
1 (bases 1 to 350)
TITLE Bombyx mori cDNA (Mita,K. 2003)
JOURNAL Unpublished (2003)
COMMENT Contact: Mita K
Genome Research Group
National Institute of Agrobiological Sciences
Owashi 1-2, Tsukuba, Ibaraki 305-8634, Japan
Email: kmita@nias.affrc.go.jp
method:uni-directional, sequence direction:sequenced from T3 primer
(5' -> 3').
Location/Qualifiers
1.350
/organism="Bombyx mori"
/mol_type="mRNA"
/db_xref="taxon:7091"
/clone="tees0720"
/tissue_type="testis"
/dev_stage="spinning stage"
/clone_lib="tees"

ORIGIN

Query Match 48.7% Score 29.2; DB 5; Length 350;
 Best Local Similarity 74.0%; Pred. No. 22;
 Matches 37; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

11 CACCAACCGGACATGCGCTAATCATGACGACCTGATACAGACCTGACACT 60
 274 CAGACCGGACACCTGAAACACGACGACGACCTGAAACGACGACCGGACACT 323

RESULT 5
 CK541226 357 bp mRNA linear EST 15-JAN-2004
 LOCUS rwbh0.009161.y1 swm Bombyx mori cDNA, mRNA sequence.
 ACCESSION CK541226
 VERSION CK541226.1 GI:40925680
 KEYWORDS EST.
 SOURCE Bombyx mori (domestic silkworm)
 ORGANISM Bombyx mori (domestic silkworm)
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Bombycoidea; Bombycidae; Bombyx.

REFERENCE
 AUTHORS Xia, Q., Zhou, Z., Lu, C., Cheng, D., Dai, F., Li, B., Zhao, P., Zha, X., Wang, Y., Zhu, Y., Yu, M., Li, C., Shen, Y., Lan, X., Yuan, L., Yang, G., Li, R., Li, H., Li, G., Su, J., Wang, X., Li, G., Zhang, Z., Li, J., Zhang, Q., Wei, N., Xu, J., Sun, H., Dong, L., Liu, D., Zhao, S., Zhao, X., Meng, Q., Lan, F., Fang, L., Li, Y., Wang, J., Ye, J., Wong, G. K.-S., Huang, H., Shi, J., Wu, Q., Li, C., Li, D., Sun, Y., Zhang, Z., Yang, Z., Huang, Y., Xi, Y., Qi, Q., He, D., Huang, H., Zhang, X., Wang, Z., Li, W. and Cao, Y.
 A draft sequence for the genome of the domesticated silkworm (Bombyx mori)
 Unpublished (2004)
 Contact: Yang Huan Ming
 Beijing Genomics Institute
 Chinese Academy of Sciences
 Beijing Airport Industrial Zone B6, Beijing 101300, P.R.China
 Email: hyang@igcp.ac.cn
 Location/Qualifiers
 1..357
 /organism="Bombyx mori"
 /mol_type="mRNA"
 /strain="Dazhao (P50)"
 /db_xref="taxon:7091"
 /sex="female"
 /tissue_type="ovary"
 /dev_stage="5ch-instar day-3 larva"
 /clone_lib="swm"
 /note="Vector: pBluescript II SK(+)"

ORIGIN
 TITLE A draft sequence for the genome of the domesticated silkworm (Bombyx mori)
 JOURNAL Unpublished (2004)
 COMMENT Contact: Yang Huan Ming
 Beijing Genomics Institute
 Chinese Academy of Sciences
 Beijing Airport Industrial Zone B6, Beijing 101300, P.R.China
 Email: hyang@igcp.ac.cn
 Location/Qualifiers
 1..357
 /organism="Bombyx mori"
 /mol_type="mRNA"
 /strain="Dazhao (P50)"
 /db_xref="taxon:7091"
 /sex="female"
 /tissue_type="ovary"
 /dev_stage="5ch-instar day-3 larva"
 /clone_lib="swm"
 /note="Vector: pBluescript II SK(+)"

FEATURES
 source

Query Match 48.7% Score 29.2; DB 7; Length 357;
 Best Local Similarity 74.0%; Pred. No. 22;
 Matches 37; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

11 CACCAACCGGACATGCGCTAATCATGACGACCTGATACAGACCTGACACT 60
 269 CAGACCGGACACCTGAAACACGACGACGACCTGAAACGACGACCGGACACT 318

RESULT 6
 CK542960 366 bp mRNA linear EST 15-JAN-2004
 LOCUS rwbh0.009514.y1 swm Bombyx mori cDNA, mRNA sequence.
 ACCESSION CK542960
 VERSION CK542960.1 GI:40927414
 KEYWORDS EST.
 SOURCE Bombyx mori (domestic silkworm)
 ORGANISM Bombyx mori (domestic silkworm)
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;

REFERENCE
 AUTHORS Xia, Q., Zhou, Z., Lu, C., Cheng, D., Dai, F., Li, B., Zhao, P., Zha, X., Wang, Y., Zhu, Y., Yu, M., Li, C., Shen, Y., Lan, X., Yuan, L., Yang, G., Li, R., Li, H., Li, G., Su, J., Wang, X., Li, G., Zhang, Z., Li, J., Zhang, Q., Wei, N., Xu, J., Sun, H., Dong, L., Liu, D., Zhao, S., Zhao, X., Meng, Q., Lan, F., Fang, L., Li, Y., Wang, J., Ye, J., Wong, G. K.-S., Huang, H., Shi, J., Wu, Q., Li, C., Li, D., Sun, Y., Zhang, Z., Yang, Z., Huang, Y., Xi, Y., Qi, Q., He, D., Huang, H., Zhang, X., Wang, Z., Li, W. and Cao, Y.
 A draft sequence for the genome of the domesticated silkworm (Bombyx mori)
 Unpublished (2004)
 Contact: Yang Huan Ming
 Beijing Genomics Institute
 Chinese Academy of Sciences
 Beijing Airport Industrial Zone B6, Beijing 101300, P.R.China
 Email: hyang@igcp.ac.cn
 Location/Qualifiers
 1..366
 /organism="Bombyx mori"
 /mol_type="mRNA"
 /strain="Dazhao (P50)"
 /db_xref="taxon:7091"
 /sex="female"
 /tissue_type="ovary"
 /dev_stage="5ch-instar day-3 larva"
 /clone_lib="swm"
 /note="Vector: pBluescript II SK(+)"

ORIGIN
 TITLE A draft sequence for the genome of the domesticated silkworm (Bombyx mori)
 JOURNAL Unpublished (2004)
 COMMENT Contact: Yang Huan Ming
 Beijing Genomics Institute
 Chinese Academy of Sciences
 Beijing Airport Industrial Zone B6, Beijing 101300, P.R.China
 Email: hyang@igcp.ac.cn
 Location/Qualifiers
 1..366
 /organism="Bombyx mori"
 /mol_type="mRNA"
 /strain="Dazhao (P50)"
 /db_xref="taxon:7091"
 /sex="female"
 /tissue_type="ovary"
 /dev_stage="5ch-instar day-3 larva"
 /clone_lib="swm"
 /note="Vector: pBluescript II SK(+)"

FEATURES
 source

Query Match 48.7% Score 29.2; DB 7; Length 366;
 Best Local Similarity 74.0%; Pred. No. 22;
 Matches 37; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

11 CACCAACCGGACATGCGCTAATCATGACGACCTGATACAGACCTGACACT 60
 269 CAGACCGGACACCTGAAACACGACGACGACCTGAAACGACGACCGGACACT 318

RESULT 7
 BP127555 370 bp mRNA linear EST 16-MAY-2003
 LOCUS BP127555 tss5 Bombyx mori cDNA clone tss50695, mRNA sequence.
 ACCESSION BP127555
 VERSION BP127555.1 GI:29560608
 KEYWORDS EST.
 SOURCE Bombyx mori (domestic silkworm)
 ORGANISM Bombyx mori (domestic silkworm)
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Bombycoidea; Bombycidae; Bombyx.

REFERENCE
 AUTHORS Wita, K., Morimoto, M., Shimada, T., Okano, K. and Maeda, S.
 TITLE Unpublished (2003)
 JOURNAL Contact: Mita K
 Genome Research Group
 National Institute of Agrobiological Sciences
 Owashi 1-2, Tsukuba, Ibaraki 305-8634, Japan
 Email: kmita@nias.affrc.go.jp
 method: uni-directional, sequence direction: sequenced from T3 primer (5' -> 3')..

FEATURES
 source
 1..370
 /organism="Bombyx mori"
 /mol_type="mRNA"
 /db_xref="taxon:7091"
 /clone="tss50695"
 /tissue_type="testis"
 /dev_stage="spinning stage"

ORIGIN /clone_1ib="tegs"

Query Match 48.7%; Score 29.2; DB 5; Length 370;
 Best Local Similarity 74.0%; Pred. No. 22;
 Matches 37; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 11 CACCACCGGACACCTGACCTGATACAGACCTGACCT 60
 Db 274 CAGCACCGGACCTGAACGACGACGACCTGAACGACCGGACCT 323

RESULT 8
 LOCUS CKS31267/c 375 bp mRNA linear EST 15-JAN-2004
 DEFINITION rswab0 009036.y1 swg Bombyx mori cDNA, mRNA sequence.
 ACCESSION CKS31267
 VERSION CKS31267.1 GI:40915720
 KEYWORDS EST.
 SOURCE Bombyx mori (domestic silkworm)
 ORGANISM Bombyx mori
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
 Bombycoidea; Bombycidae; Bombyx.

REFERENCE 1 (bases 1 to 375)
 Xia, Q., Zhou, Z., Lu, C., Cheng, D., Dai, F., Li, B., Zhao, P., Zha, X.,
 Cheng, T., Chai, C., Pan, G., Xu, J., Liu, C., Lin, Y., Qian, J., Hou, Y.,
 Wu, Z., Li, G., Pan, M., Li, C., Shen, Y., Lan, X., Yuan, L., Yang, G.,
 Wan, Y., Zhu, Y., Yu, M., Shen, W., Wu, D., Xiang, Z., Yu, J., Wang, J.,
 Li, R., Li, H., Li, G., Su, J., Wang, X., Li, G., Zhang, Z., Li, J.,
 Zhang, Q., Wei, N., Xu, J., Sun, H., Dong, L., Liu, D., Zhao, S., Zhao, X.,
 Meng, Q., Lan, F., Fang, L., Li, Y., Wang, J., Ye, J., Wong, G. K.-S.,
 Yang, H., Shi, J., Wu, Q., Li, C., Li, D., Sun, Y., Zhang, Z., Yang, Z.,
 Huang, Y., Xi, Y., Qi, Q., He, D., Huang, H., Zhang, X., Wang, Z., Li, W.
 and Cao, Y.
 A draft sequence for the genome of the domesticated silkworm
 (Bombyx mori)

JOURNAL Unpublished (2004)
 COMMENT Contact: Yang Huan Ming
 Beijing Genomics Institute
 Chinese Academy of Sciences
 Beijing Airport Industrial Zone B6, Beijing 101300, P.R.China
 Email: hyang@igtp.ac.cn

FEATURES
 source location/Qualifiers
 1..375
 /organism="Bombyx mori"
 /mol_type="mRNA"
 /strain="Dazhao (P50)"
 /db_xref="taxon:7091"
 /sex="male"
 /tissue_type="testis"
 /dev_stage="5th-instar day-3 larva"
 /clone_1ib="swg"
 /note="Vector: pBluescript II SK(+)"

ORIGIN

Query Match 48.7%; Score 29.2; DB 7; Length 375;
 Best Local Similarity 74.0%; Pred. No. 22;
 Matches 37; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 11 CACCACCGGACACCTGACCTGATACAGACCTGACCT 60
 Db 267 CAGCACCGGACCTGAACGACGACGACCTGAACGACCGGACCT 218

RESULT 9
 LOCUS CKS42714 375 bp mRNA linear EST 15-JAN-2004
 DEFINITION rswab0 008394.y1 swb Bombyx mori cDNA, mRNA sequence.
 ACCESSION CKS42714
 VERSION CKS42714.1 GI:40927168
 KEYWORDS EST.
 SOURCE Bombyx mori (domestic silkworm)

ORGANISM Bombyx mori
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
 Bombycoidea; Bombycidae; Bombyx.

REFERENCE 1 (bases 1 to 375)
 Xia, Q., Zhou, Z., Lu, C., Cheng, D., Dai, F., Li, B., Zhao, P., Zha, X.,
 Cheng, T., Chai, C., Pan, G., Xu, J., Liu, C., Lin, Y., Qian, J., Hou, Y.,
 Wu, Z., Li, G., Pan, M., Li, C., Shen, Y., Lan, X., Yuan, L., Yang, G.,
 Wan, Y., Zhu, Y., Yu, M., Shen, W., Wu, D., Xiang, Z., Yu, J., Wang, J.,
 Li, R., Li, H., Li, G., Su, J., Wang, X., Li, G., Zhang, Z., Li, J.,
 Zhang, Q., Wei, N., Xu, J., Sun, H., Dong, L., Liu, D., Zhao, S., Zhao, X.,
 Meng, Q., Lan, F., Fang, L., Li, Y., Wang, J., Ye, J., Wong, G. K.-S.,
 Yang, H., Shi, J., Wu, Q., Li, C., Li, D., Sun, Y., Zhang, Z., Yang, Z.,
 Huang, Y., Xi, Y., Qi, Q., He, D., Huang, H., Zhang, X., Wang, Z., Li, W.
 and Cao, Y.
 A draft sequence for the genome of the domesticated silkworm
 (Bombyx mori)

JOURNAL Unpublished (2004)
 COMMENT Contact: Yang Huan Ming
 Beijing Genomics Institute

FEATURES
 source location/Qualifiers
 1..375
 /organism="Bombyx mori"
 /mol_type="mRNA"
 /strain="Dazhao (P50)"
 /db_xref="taxon:7091"
 /sex="female"
 /tissue_type="ovary"
 /dev_stage="5th-instar day-3 larva"
 /clone_1ib="swb"
 /note="Vector: pBluescript II SK(+)"

ORIGIN

Query Match 48.7%; Score 29.2; DB 7; Length 375;
 Best Local Similarity 74.0%; Pred. No. 22;
 Matches 37; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 11 CACCACCGGACACCTGACCTGATACAGACCTGACCT 60
 Db 278 CAGCACCGGACCTGAACGACGACGACCTGAACGACCGGACCT 327

RESULT 10
 LOCUS CK487724 378 bp mRNA linear EST 14-JAN-2004
 DEFINITION rswab0 005006.y1 swa Bombyx mori cDNA, mRNA sequence.
 ACCESSION CK487724
 VERSION CK487724.1 GI:40843903
 KEYWORDS EST.
 SOURCE Bombyx mori (domestic silkworm)
 ORGANISM Bombyx mori
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
 Bombycoidea; Bombycidae; Bombyx.

REFERENCE 1 (bases 1 to 378)
 Xia, Q., Zhou, Z., Lu, C., Cheng, D., Dai, F., Li, B., Zhao, P., Zha, X.,
 Cheng, T., Chai, C., Pan, G., Xu, J., Liu, C., Lin, Y., Qian, J., Hou, Y.,
 Wu, Z., Li, G., Pan, M., Li, C., Shen, Y., Lan, X., Yuan, L., Yang, G.,
 Wan, Y., Zhu, Y., Yu, M., Shen, W., Wu, D., Xiang, Z., Yu, J., Wang, J.,
 Li, R., Li, H., Li, G., Su, J., Wang, X., Li, G., Zhang, Z., Li, J.,
 Zhang, Q., Wei, N., Xu, J., Sun, H., Dong, L., Liu, D., Zhao, S., Zhao, X.,
 Meng, Q., Lan, F., Fang, L., Li, Y., Wang, J., Ye, J., Wong, G. K.-S.,
 Yang, H., Shi, J., Wu, Q., Li, C., Li, D., Sun, Y., Zhang, Z., Yang, Z.,
 Huang, Y., Xi, Y., Qi, Q., He, D., Huang, H., Zhang, X., Wang, Z., Li, W.
 and Cao, Y.
 A draft sequence for the genome of the domesticated silkworm
 (Bombyx mori)

JOURNAL Unpublished (2004)
 COMMENT Contact: Yang Huan Ming
 Beijing Genomics Institute

TITLE	A draft sequence for the genome of the domesticated silkworm (Bombyx mori)			
JOURNAL	Unpublished (2004)			
COMMENT	Contact: Yang Huan Ming Beijing Genomics Institute Chinese Academy of Sciences Beijing Airport Industrial Zone B6, Beijing 101300, P.R.China Email: hyang@igtp.ac.cn.			
FEATURES	Location/Qualifiers			
source	1..400			
	/organism="Bombyx mori"			
	/mol_type="mRNA"			
	/strain="DaZhao (P50)"			
	/db_xref="taxon:7091"			
	/sex="female"			
	/tissue_type="ovary"			
	/dev_stage="5th-instar day-3 larva"			
	/clone_id="BwH"			
	/note="Vector: pBluescript II SK(+)"			
ORIGIN				
Query Match	48.7%	Score 29.2;	DB 7;	Length 400;
Best Local Similarity	74.0%;	Pred. No. 22;		
Matches	37;	Conservative 0;	Mismatches 13;	Indels 0;
				Gaps 0;
OR	11	CACCAACCGGACATGCGCTGAACATAGACAGCTGATACAAGACCTGCACCT	60	
Dp	217	CAGCAGCGGACGACCTGACACGACAGCAGACCTGACACGACCGGACCT	168	
RESULT 14				
LOCUS	CK501702.c			
DEFINITION	rBwCC0_001340.y1 swc Bombyx mori cDNA, mRNA sequence.			
ACCESSION	CK501702			
VERSION	CK501702.1 GI:40862441			
KEYWORDS	EST.			
SOURCE	Bombyx mori (domestic silkworm)			
ORGANISM	Bombyx mori			
REFERENCE	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Bombycoidea; Bombycidae; Bombyx.			
AUTHORS	1 (bases 1 to 407)			
	Xie,Q., Zhou,Z., Lu,C., Cheng,D., Bai,F., Li,B., Zhao,P., Zha,X., Cheng,Y., Chai,C., Pan,G., Xu,J., Liu,C., Lin,Y., Qian,J., Hou,Y., Wu,Z., Li,G., Pan,M., Li,C., Shen,Y., Lan,X., Yuan,L., Yang,G., Wan,Y., Zhu,Y., Yu,M., Shen,W., Wu,D., Xiang,Z., Yu,J., Wang,J., Li,R., Li,H., Li,G., Su,J., Wang,X., Li,G., Zhang,Z., Li,J., Zhang,Q., Wei,N., Xu,J., Sun,H., Dong,L., Liu,D., Zhao,S., Zhao,X., Meng,Q., lan,F., Fang,L., Li,Y., Wang,D., Ye,J., Wong,G.K.-S., Yang,H., Shi,J., Wu,Q., Li,C., Li,D., Sun,Y., Zhang,Z., Yang,Z., Huang,Y., Xi,Y., Qi,Q., He,D., Huang,H., Zhang,X., Wang,Z., Li,W. and Cao,Y.			
TITLE	A draft sequence for the genome of the domesticated silkworm (Bombyx mori)			
JOURNAL	Unpublished (2004)			
COMMENT	Contact: Yang Huan Ming Beijing Genomics Institute Chinese Academy of Sciences Beijing Airport Industrial Zone B6, Beijing 101300, P.R.China Email: hyang@igtp.ac.cn.			
FEATURES	Location/Qualifiers			
source	1..407			
	/organism="Bombyx mori"			
	/mol_type="mRNA"			

```

/strain="Dazhao (P50) "
/db_xref="taxon:7091"
/sex="female"
/tissue_type="Fat body"
/dev_stage="5th-instar day-3 larva"
/clone_lib="swc"
/note="Vector: pBluescript II SK(+)"

ORIGIN
Query Match      48.7%; Score 29.2; DB 7; Length 407;
Best Local Similarity 74.0%; Pred. No. 22;
Matches 37; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy
11 CACCACCGGCACATGCGCTTAACATAGACCTGTATACAAGACCTGCACCT 60
|||||
382 CAGCACCGGCACCTGAACACGACGACGACCTGAACACGACCGCACCT 333

RESULT 15
LOCUS          CK488576/c              417 bp    mRNA    linear    EST 14-JAN-2004
DEFINITION     rswb0.006166.y1 swa Bombyx mori cDNA, mRNA sequence.
ACCESSION      CK488576
VERSION        CK488576.1 GI:40844755
KEYWORDS
SOURCE         Bombyx mori (domestic silkworm)
ORGANISM       Bombyx mori
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
Bombycoidea; Bombycidae; Bombyx.
1 (bases 1 to 417)
Xia,Q., Zhou,Z., Lu,C., Cheng,D., Dai,F., Li,B., Zhao,P., Zha,X.,
Cheng,T., Chai,C., Pan,G., Xu,J., Liu,C., Lin,Y., Qian,J., Hou,Y.,
Wu,Z., Li,G., Pan,M., Li,C., Shen,Y., Lan,X., Yuan,L., Yang,G.,
Wan,Y., Zhu,Y., Yu,M., Shen,W., Wu,D., Xiang,Z., Yu,J., Wang,J.,
Li,R., Li,H., Li,G., Su,J., Wang,X., Li,G., Zhang,Z., Li,J.,
Zhang,Q., Wei,N., Xu,J., Sun,H., Dong,L., Liu,D., Zhao,S., Zhao,X.,
Meng,Q., Lan,F., Fang,L., Li,Y., Wang,J., Ye,J., Wong,G.K.-S.,
Yang,H., Shi,Y., Wu,Q., Li,C., Li,D., Sun,Y., Zhang,Z., Yang,Z.,
Huang,Y., Xi,Y., Qi,Q., He,D., Huang,H., Zhang,X., Wang,Z., Li,W.
and Cao,Y.
A draft sequence for the genome of the domesticated silkworm
(Bombyx mori)
Unpublished (2004)
Contact: Yang Huan Ming
Beijing Genomics Institute
Chinese Academy of Sciences
Beijing Airport Industrial Zone B6, Beijing 101300, P.R.China
Email: hyanmg@igpp.ac.cn.
Location/Qualifiers
1..417
/organism="Bombyx mori"
/mol_type="mRNA"
/strain="Dazhao (P50) "
/db_xref="taxon:7091"
/sex="mixed"
/tissue_type="silk gland"
/dev_stage="5th-instar day-3 larva"
/clone_lib="swa"
/note="Vector: pBluescript II SK(+)"

ORIGIN
Query Match      48.7%; Score 29.2; DB 7; Length 417;
Best Local Similarity 74.0%; Pred. No. 22;
Matches 37; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy
11 CACCACCGGCACATGCGCTTAACATAGACCTGTATACAAGACCTGCACCT 60
|||||
Db      149 CAGCACCGGCACCTGAACACGACGACGACCTGAACACGACCGCACCT 100

```


This Page Blank (uspto)


```
XX
SQ Sequence 60 BP; 17 A; 21 C; 13 G; 9 T; 0 U; 0 Other;
XX
Query Match 100.0%; Score 60; DB 2; Length 60;
Best Local Similarity 100.0%; Pred. No. 2.6e-11;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCACTACTGCAACCGGCGCATGGCTTAACATCAGACCTGATACAGACCTGACCT 60
DB 1 GGCACTACTGCAACCGGCGCATGGCTTAACATCAGACCTGATACAGACCTGACCT 60
XX
RESULT 2
AAV48325
ID AAV48325 standard; DNA; 60 BP.
XX
AC AAV48325;
XX
DT 20-NOV-1998 (first entry)
XX
DE Nucleotide sequence encoding MUC1 tandem repeat unit R10.
XX
KM 89; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
KM tumour; tumour-associated antigen.
XX
OS Homo sapiens.
XX
MO9837095-A2.
XX
PN 27-AUG-1998.
XX
PD 24-FEB-1998; 98WO-US003693.
XX
PF 24-FEB-1997; 97US-0038253P.
XX
PR 24-FEB-1997; 97US-0038253P.
XX
XX
PA (THER-) THERION BIOLOGICS CORP.
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA (DAND) DANA FARBER CANCER INST INC.
XX
PI Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;
XX
DR WPI; 1998-467492/40.
XX
XX
PT New recombinant pox virus for tumour therapy - comprises DNA encoding an
PT immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
XX
PS Disclosure; Page 11; 42pp; English.
XX
XX
CC The MUC1 tandem repeat units AAV48317-V48325 were used to create an
CC immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
CC (RPV). The RPV was used in a pharmaceutical composition also containing
CC an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The
CC recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
CC does not undergo significant genetic deletion, thereby providing an
CC unexpectedly stable and immunogenic pox virus. They can be used to
CC prevent or treat tumours expressing MUC1 tumour-associated antigens
XX
SQ Sequence 60 BP; 12 A; 22 C; 19 G; 7 T; 0 U; 0 Other;
XX
Query Match 62.7%; Score 37.6; DB 2; Length 60;
Best Local Similarity 76.7%; Pred. No. 0.0016;
Matches 46; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 1 GGCACTACTGCAACCGGCGCATGGCTTAACATCAGACCTGATACAGACCTGACCT 60
DB 1 GGTAATAACAGCGCCACCCGACATGGCTTCAAGAGCGCTCCGATACAGACCGCGCT 60
XX
RESULT 3
AAD00385
ID AAD00385 standard; DNA; 525 BP.
XX
AC AAD00385;
```

```
XX
DT 29-AUG-2000 (first entry)
XX
DE Human Mucin 1 (MUC-1) protein fragment encoding DNA #2.
XX
KM Human Mucin 1; MUC-1; tumour; pMRS30 expression vector; anti-tumour;
KM therapy; immune response; cytostatic; vaccine; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..525
FT /*tag= a
FT /product= "MUC-1 protein fragment"
XX
MO200025827-A2.
XX
PN 11-MAY-2000.
XX
PD 18-OCT-1999; 99WO-EP007874.
XX
PF 30-OCT-1998; 98IT-MI002330.
XX
PR (MENA) MENARINI RICERCHE SPA.
XX
PA Parente D, Di Massimo AM, De Santis R;
XX
PI WPI; 2000-365410/31.
XX
DR P-PSDB; AAY71021.
XX
XX
PT Composition containing one or more DNA molecules encoding fragments of a
PT Mucin 1 (MUC-1) protein overexpressed in tumor cells, useful in anti-
PT tumor therapy.
XX
PS Claim 16; Fig 2; 56pp; English.
XX
XX
CC The present sequence is a DNA encoding a fragment of human Mucin 1 (MUC-
CC 1) antigenic protein which is overexpressed in tumour cells. The sequence
CC was obtained from B120 tumour cells by reverse transcriptase-PCR and
CC corresponds to nucleotides 205-720 of the EMBL sequence J05581 with a
CC start codon and two stop codons. The present sequence is cloned into a
CC pMRS30 expression vector and used in pharmaceutical composition e.g.
CC vaccine for inducing an antigen-specific anti-tumour immune response.
CC Composition containing this DNA molecule is useful in anti-tumour therapy
CC of patients affected with tumours characterised by high MUC-1 expression
XX
SQ Sequence 525 BP; 111 A; 199 C; 128 G; 87 T; 0 U; 0 Other;
XX
Query Match 61.0%; Score 36.6; DB 3; Length 525;
Best Local Similarity 76.3%; Pred. No. 0.0061;
Matches 45; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 1 GGCACTACTGCAACCGGCGCATGGCTTAACATCAGACCTGATACAGACCTGACCT 59
DB 256 GGAAGTACTGCTCTCCACGACGACGAGGTGTACTCTCGGCTCCGATACAGAGCGCGCCC 314
XX
RESULT 4
AAD00391
ID AAD00391 standard; DNA; 891 BP.
XX
AC AAD00391;
XX
DT 15-SEP-2003 (revised)
DT 29-AUG-2000 (first entry)
XX
XX
DE Ubiquitin-E. coli LacI-human Mucin 1 fusion protein encoding DNA #2.
XX
KM Ubiquitin; LacI; beta-galactosidase; fusion protein; human; Mucin 1;
KM MUC-1; tumour; pMRS30 expression vector; anti-tumour; therapy;
KM immune response; cytostatic; vaccine; ds.
XX
OS Homo sapiens.
```

OS		Bescherichia coli.	
OS		Chimeric.	
XX			
FH	Key	Location/Qualifiers	
FT	CDS	1..891	
FT		/tag= a	
FT		/product= "Ubiquitin-LacI-MUC-1 fusion protein"	
FT	misc_feature	1..369	
FT		/tag= b	
FT		/label= UBILACI_DNA	
FT	misc_feature	/note= "Includes ubiquitin-E. coli LacI fusion DNA"	
FT		370..891	
FT		/tag= c	
FT		/note= "Human MUC-1 partial DNA that corresponds to nucleotides 205-720 of the EMBL sequence J05581 with two stop codons"	
FT			
PN	WO20025827-A2.		
PD	11-MAY-2000.		
PF	18-OCT-1999;	99WO-BP007874.	
PR	30-OCT-1998;	98IT-WI002330.	
PA	(MENA) MENARINI RICERCHE SPA.		
PI	Parente D, Di Massimo AM, De Santis R;		
DR	WPI; 2000-365410/31.		
DR	P-PsDB; AAU71027.		
PT	Composition containing one or more DNA molecules encoding fragments of a		
PT	Mucin 1 (MUC-1) protein overexpressed in tumor cells, useful in anti-		
PT	tumor therapy.		
PS	Claim 18; Fig 8; 56pp; English.		
CC	The present sequence is a DNA encoding a fusion protein consisting of		
CC	human Mucin 1 (MUC-1) fragment fused to Ubilaci sequence at the N-		
CC	terminus. The Ubilaci segment consists of ubiquitin from MCP7 cell line		
CC	and a portion of E. coli beta-galactosidase (lacI). MUC-1 is an antigenic		
CC	protein overexpressed in tumour cells. The present sequence is cloned		
CC	into a pMS30 expression vector and used in pharmaceutical composition		
CC	e.g., vaccine for inducing an antigen-specific anti-tumour immune		
CC	response. Composition containing this DNA molecule is useful in anti-		
CC	tumour therapy of patients affected with tumours characterised by high		
CC	MUC-1 expression. (Updated on 15-Sep-2003 to standardise OS field)		
SQ	Sequence 891 BP; 203 A; 307 C; 225 G; 156 T; 0 U; 0 Other;		
Query Match	61.0%; Score 36.6; DB 3; Length 891;		
Best Local Similarity	76.3%; Pred No. 0 0069;		
Matches	45; Conservative 0; Mismatches 14; Indels 0; Gaps 0		
Dy	1 GGCACTACTGTCACCGAGCATGTGGCTATGCAGCACCTGATACAAGAAGCTTGCACC 59 622 GGAATTACTTGCTCCTCCACGACGACGAGTGTTAATCTGCGGCTCCGAGATACAGCGCGGCC 680 		
RESULT 5			
AADV48323			
ID	AAV48323 standard; DNA; 60 BP.		
XX			
AC	AAV48323;		
DT	20-NOV-1998 (first entry)		
Nucleotide sequence encoding MUC1 tandem repeat unit R8.			
DE			
XX			
KM	ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen; tumour; tumour-associated antigen.		
XX			

OS	Hom	sapiens.
PN	WO9837095-A2.	
PD	27-AUG-1998.	
PF	24-FEB-1998;	98WO-US003693.
PR	24-FEB-1997;	97US-0038253P.
PA	(THER-) THERION BIOLOGICS CORP.	
PA	(USSH) US DEPT HEALTH & HUMAN SERVICES.	
PA	(DAND) DANA FARNER CANCER INST INC.	
PI	Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;	
DR	WPI, 1998-467492/40.	
PT	New recombinant pox virus for tumour therapy - comprises DNA encoding an immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.	
PS	Disclosure; Page 11; 42pp; English.	
CC	The MUC1 tandem repeat units AAV48317-V48325 were used to create an immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus (RPV). The RPV was used in a pharmaceutical composition also containing an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The recombinant pox virus therefore encodes an immunogenic MUC1 fragment that does not undergo significant genetic deletion, thereby providing an unexpectedly stable and immunogenic pox virus. They can be used to prevent or treat tumours expressing MUC1 tumour-associated antigens	
SQ	Sequence 60 BP; 10 A; 24 C; 17 G; 9 T; 0 U; 0 Other;	
Qy	Query Match 59.0%; Score 35.4; DB 2; Length 60;	
Db	Best Local Similarity 79.2%; Pred.No. 0.0095;	
	Matches 42; Conservative 0; Mismatches 11; Indels 0; Gaps 0	
	7 ACTGCACCGACGGGACATGGCGTAACATCAGCACCTGTATACAGACTGCACC 59	
	7 ACTGCCTCTCGGGCGCATGTGTGACCTCAGCTCTGTACACAGCGAGCCCC 59	
RESULT 6		
AAD00388		
ID	AAD00388 standard; DNA; 1371 BP.	
AC	AAD00388;	
DT	29-AUG-2000 (first entry)	
DE	Human Mucin 1 (MUC-1) protein fragment encoding DNA #5.	
KM	Human; Mucin 1; MUC-1; tumour; pKRS30 expression vector; anti-tumour; therapy; immune response; cytostatic; vaccine; ds.	
OS	Hom	sapiens.
PH	Key	Location/Qualifiers
FT	CDS	1..1371
FT		/*tag= a
PN		/product= "MUC-1 protein fragment"
PD	WO200025827-A2.	
PF	11-MAY-2000.	
PF	18-OCT-1999;	99WO-EP007874.
PR	30-OCT-1998;	98IT-MI002330.
PA	(MENA) MENARINI RICERCHE SPA.	

```

PI Parente D, Di Massimo AM, De Santis R;
XX WPI; 2000-365410/31.
DR P-PSDB; AAY71024.
XX
XX Composition containing one or more DNA molecules encoding fragments of a
PT Mucin 1 (MUC-1) protein overexpressed in tumor cells, useful in anti-
PT tumor therapy.
XX
XX Claim 16; Fig 5; 56pp; English.
XX
XX The present sequence is a DNA encoding a fragment of human Mucin 1 (MUC-
CC 1) antigenic protein which is overexpressed in tumor cells. The sequence
CC was obtained by PCR from plasmids pMR5166, pMR5167, pMR5168 and pMR5169
CC which contain MUC-1 DNA from BT20 tumour cells. It corresponds to
CC nucleotides 136-1497 of the EMBL sequence J05581 with a start codon and
CC two stop codons. The present sequence is cloned into a pMR30 expression
CC vector and used in pharmaceutical composition e.g. Vaccine for inducing
CC an antigen-specific anti-tumour immune response. Composition containing
CC this DNA molecule is useful in anti-tumour therapy of patients affected
CC with tumours characterised by high MUC-1 expression
SQ
SQ Sequence 1371 BP; 300 A; 450 C; 321 G; 300 T; 0 U; 0 Other;
Query Match 58.3%; Score 35; DB 3; Length 1371;
Best Local Similarity 74.6%; Pred. No. 0.027;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 1 GGCAGTCTGCACACCGGACATGCGGTACATCAGACCTGTATCAAGACCTGCACC 59
DB 325 GGAAGTACCGCTCCACGACGACACGGGTATCCTCGGCTCCGATACCAAGCGGCCCCC 383
RESULT 7
AAD00394
ID AAD00394 standard; DNA; 1737 BP.
XX
XX AAD00394;
XX
XX 15-SEP-2003 (revised)
DT 29-AUG-2000 (first entry)
XX
DE Ubiquitin-E. coli LacI-human Mucin 1 fusion protein encoding DNA #5.
XX
XX Ubiquitin; LacI; beta-galactosidase; fusion protein; human; Mucin 1;
KW MUC-1; tumour; pMR30 expression vector; anti-tumour; therapy;
XX immune response; cytosolic; vaccine; ds.
XX
XX Homo sapiens.
OS Escherichia coli.
OS Chimeric.
XX
XX Key Location/Qualifiers
FH 1..1737
FT CDS
FT /tag= a
FT /product= "Ubiquitin-LacI-MUC-1 fusion protein"
FT /tag= b
FT /label= UBILacI DNA
FT /note= "Includes ubiquitin-E. coli LacI fusion DNA"
FT 370..1737
FT /tag= C
FT /note= "Human MUC-1 partial DNA that corresponds to
FT nucleotides 136-1497 of the EMBL sequence J05581 with two
FT stop codons"
XX
XX WO200025827-A2.
XX
XX 11-MAY-2000.
XX
XX 18-OCT-1999; 99WO-EP007874.
XX
XX 30-OCT-1998; 98IT-MI002330.
XX
XX PR

```

```

XX
XX (MENA ) MENARINI RICERCH SPA.
PA Parente D, Di Massimo AM, De Santis R;
XX WPI; 2000-365410/31.
DR P-PSDB; AAY71030.
XX
XX Composition containing one or more DNA molecules encoding fragments of a
PT Mucin 1 (MUC-1) protein overexpressed in tumor cells, useful in anti-
PT tumor therapy.
XX
XX Claim 18; Fig 11; 56pp; English.
XX
XX The present sequence is a DNA encoding a fusion protein consisting of
CC human Mucin 1 (MUC-1) fragment fused to UBILacI sequence at the N-
CC terminus. The UBILacI sequence consists of ubiquitin from MCF7 cell line
CC and a portion of E. coli beta-galactosidase (LacI). MUC-1 is an antigenic
CC protein overexpressed in tumour cells. The present sequence is cloned
CC into a pMR30 expression vector and used in pharmaceutical composition
CC e.g. vaccine for inducing an antigen-specific anti-tumour immune
CC response. Composition containing this DNA molecule is useful in anti-
CC tumour therapy of patients affected with tumours characterised by high
CC MUC-1 expression. (updated on 15-SEP-2003 to standardise OS field)
SQ
SQ Sequence 1737 BP; 392 A; 558 C; 418 G; 369 T; 0 U; 0 Other;
Query Match 58.3%; Score 35; DB 3; Length 1737;
Best Local Similarity 74.6%; Pred. No. 0.029;
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 1 GGCAGTCTGCACACCGGACATGCGGTACATCAGACCTGTATCAAGACCTGCACC 59
DB 691 GGAAGTACCGCTCCACGACGACACGGGTATCCTCGGCTCCGATACCAAGCGGCCCCC 749
RESULT 8
AAV48321
ID AAV48321 standard; DNA; 60 BP.
XX
XX AAV48321;
XX
XX 20-NOV-1998 (first entry)
DT
XX
DE Nucleotide sequence encoding MUC1 tandem repeat unit R6.
XX
XX ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
KW tumour; tumour-associated antigen.
XX
XX Homo sapiens.
OS
XX
XX WO9837095-A2.
XX
XX 27-AUG-1998.
XX
XX 24-FEB-1998; 98WO-US003693.
XX
XX 24-FEB-1997; 97US-0038253P.
XX
XX (THER-) THERION BIOLOGICS CORP.
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX (DAND ) DANA FARBER CANCER INST INC.
XX
XX Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;
XX WPI; 1998-467492/40.
XX
XX New recombinant pox virus for tumour therapy - comprises DNA encoding an
XX immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
XX Disclosure; Page 11; 42pp; English.
XX
XX The MUC1 tandem repeat units AAV48317-V48325 were used to create an

```

CC immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
 CC (RPV). The RPV was used in a pharmaceutical composition also containing
 CC an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The
 CC recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
 CC does not undergo significant genetic deletion, thereby providing an
 CC unexpectedly stable and immunogenic pox virus. They can be used to
 CC prevent or treat tumours expressing MUC1 tumour-associated antigens
 XX
 SQ Sequence 60 BP; 13 A; 27 C; 17 G; 3 T; 0 U; 0 Other;

Query Match 55.7%; Score 33.4; DB 2; Length 60;
 Best Local Similarity 72.9%; Pred. No. 0.048;
 Matches 43; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1 GGCAGTACTGCACCGGACATGGCGTAACTACAGCAGCCTGACCTGACCC 59
 DB 1 GGCAGCAGCGACCGCCGCCGACACGAGGCGCCAGACAGCAGCAGCCTGCGCC 59

RESULT 9

ID AAV48322 standard; DNA; 60 BP.

AC AAV48322;
 DT 20-NOV-1998 (first entry)

DE Nucleotide sequence encoding MUC1 tandem repeat unit R7.

KW ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
 KM tumour; tumour-associated antigen.

OS Homo sapiens.

PN WO937095-A2.

PD 27-AUG-1998.

PF 24-FEB-1998; 98WO-US003693.

PR 24-FEB-1997; 97US-0038253P.

PA (THER-) THERION BIOLOGICS CORP.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

PA (DAND) DANA FARBER CANCER INST INC.

PI Schlom J, Kantox J, Kufe D, Panicali D, Gritz L;

DR WPI; 1998-467492/40.

PT New recombinant pox virus for tumour therapy - comprises DNA encoding an
 PT immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.

PS Disclosure; Page 11; 42pp; English.

CC The MUC1 tandem repeat units AAV48317-V48325 were used to create an
 CC immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
 CC (RPV). The RPV was used in a pharmaceutical composition also containing
 CC an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The
 CC recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
 CC does not undergo significant genetic deletion, thereby providing an
 CC unexpectedly stable and immunogenic pox virus. They can be used to
 CC prevent or treat tumours expressing MUC1 tumour-associated antigens
 XX

SQ Sequence 60 BP; 14 A; 24 C; 16 G; 6 T; 0 U; 0 Other;

Query Match 55.7%; Score 33.4; DB 2; Length 60;
 Best Local Similarity 72.9%; Pred. No. 0.048;
 Matches 43; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1 GGCAGTACTGCACCGGACATGGCGTAACTACAGCAGCCTGACCTGACCC 59
 DB 1 GGAAGTACCGCTCAGCCTGACACGAGGCGCCAGACAGCAGCAGCCTGCGCC 59

RESULT 10

ID AAV48320 standard; DNA; 60 BP.

AC AAV48320;
 DT 20-NOV-1998 (first entry)

DE Nucleotide sequence encoding MUC1 tandem repeat unit R5.

KW ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
 KM tumour; tumour-associated antigen.

OS Homo sapiens.

PN WO937095-A2.

PD 27-AUG-1998.

PF 24-FEB-1998; 98WO-US003693.

PR 24-FEB-1997; 97US-0038253P.

PA (THER-) THERION BIOLOGICS CORP.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

PA (DAND) DANA FARBER CANCER INST INC.

PI Schlom J, Kantox J, Kufe D, Panicali D, Gritz L;

DR WPI; 1998-467492/40.

PT New recombinant pox virus for tumour therapy - comprises DNA encoding an
 PT immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.

PS Disclosure; Page 11; 42pp; English.

CC The MUC1 tandem repeat units AAV48317-V48325 were used to create an
 CC immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus
 CC (RPV). The RPV was used in a pharmaceutical composition also containing
 CC an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The
 CC recombinant pox virus therefore encodes an immunogenic MUC1 fragment that
 CC does not undergo significant genetic deletion, thereby providing an
 CC unexpectedly stable and immunogenic pox virus. They can be used to
 CC prevent or treat tumours expressing MUC1 tumour-associated antigens
 XX

SQ Sequence 60 BP; 9 A; 27 C; 15 G; 9 T; 0 U; 0 Other;

Query Match 55.3%; Score 33.2; DB 2; Length 60;
 Best Local Similarity 75.9%; Pred. No. 0.056;
 Matches 41; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 7 ACTGCACCGGACGATGCGTAACTACAGCAGCCTGACCTGACCT 60
 DB 7 ACGGCCCCCTCTCAGCGTAACTACGCCCCGATACAGACCGCCCCCT 60

RESULT 11

ID AAV48324 standard; DNA; 60 BP.

AC AAV48324;
 DT 20-NOV-1998 (first entry)

DE Nucleotide sequence encoding MUC1 tandem repeat unit R9.

KW ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
 KM tumour; tumour-associated antigen.

OS Homo sapiens.


```

Db      693   GGCTCCACCGCCCCCAGCCCAAGGATGTACCTGGGCCCCGGAGACACAGCGCGGCC 751
RESULT 14
ID      ACC55474/c
XX      ACC55474 standard; cDNA, 584 BP.
AC      ACC55474;
XX      27-JUN-2003 (first entry)
DT
DE      Rice endosperm EST #35.
XX
XX      rice; endosperm; expressed sequence tag; microarray; EST; ss.
XX
OS      Oryza sativa.
XX
PN      CN1366073-A.
XX
PD      28-AUG-2002.
XX
PF      31-OCT-2001; 2001CN-00135869.
XX
PR      31-OCT-2001; 2001CN-00135869.
XX
PA      (UYZH-) UNIV ZHEJIANG.
PI      Dong H, Li D;
DR      WPI; 2003-382548/37.
XX
PT      Rice endosperm specific expression sequence label and its gene chip.
PS      Claim 1; Page 22-23; 29pp; Chinese.
XX
CC      The invention relates to a novel kind of rice endosperm specific
CC      expressed sequence tags (EST's) and its gene chip are disclosed. The EST
CC      technique is used for DNA sequencing to configure rice endosperm cDNA
CC      library. The obtained expression sequence label goes through removing
CC      redundant sequences, indexing internet database, sorting and collecting
CC      non-redundant sequences to obtain 50 novel expression sequence labels.
CC      The gene chip is prepared from them by the microarray technique and has
CC      broad-spectrum application. The sequences shown in ACC55440-ACC55489
CC      represent the rice endosperm EST's of the invention
XX
SQ      Sequence 584 BP; 148 A; 109 C; 158 G; 169 T; 0 U; 0 Other;
Query Match          46.7%; Score 28; DB 8; Length 584;
Best Local Similarity 71.2%; Pred. No. 6.2;
Matches 37; Conservative 0; Mismatches 15; Indels 0; Gaps 0
OY      8 GTGCACCAACCGGACATCGGCTTAACATCATGACACTGATCAAGACCTGCACC 59
DB      214 CAGCACCAACCGGACGAGAGACGACGACGACAAAGCTTACACCAACCAAC 163
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
RESULT 15
ADG14257
ID      ADG14257 standard; DNA; 717 BP.
AC      ADG14257;
XX
XX      26-FEB-2004 (first entry)
DT
DE      Cellulose binding region coding sequence.
XX
XX      Microarray; cellulose binding region; cellulase; cellulose; gene; db.
XX
OS      Clostridium thermocellum.
XX
Key     Location/Qualifiers
FT      1..717
TT      /*tag= a

```

```

FT      /product= "Cellulose binding region"
PM      WO2003074722-A2.
PD      12-SEP-2003.
PF      06-MAR-2003; 2003WO-IL000177.
PR      07-MAR-2002; 2002US-0362061P.
XX      (ZEPH-) ZEPHYR PROTEOMIX LTD.
XX      PA
XX      PI      Morag E;
XX      DR      WPI; 2003-845072/78.
XX      DR      P-PSDB; ADG14258.
XX      PT      New microarray of cellulose binding chimeric proteins, useful for
XX      PT      displaying peptide libraries, random or rationally designed, and for high
XX      PT      throughput screening for ligands, epitopes or ligand binding sites.
XX      PS      Disclosure; Fig 1; 86pp; English.
XX      CC      The present invention relates to microarrays of proteins comprising a
XX      CC      cellulose binding region as a means for binding to a cellulase substrate
XX      CC      such as cellulose. The invention is useful for generating or producing an
XX      CC      array and screening a plurality of peptides or biologically active
XX      CC      moieties for interactions with one or more components of a fluid sample.
XX      CC      The microarray is useful for displaying peptide libraries, random or
XX      CC      rationally designed, and for high throughput screening for ligands,
XX      CC      epitopes or ligand binding sites. The present sequence is the coding
XX      CC      sequence for the cellulose binding region with the N-terminal and C-
XX      CC      terminal linker regions from Clostridium thermocellum, used to illustrate
XX      CC      the invention.
XX      QO      Sequence 717 BP; 256 A; 179 C; 135 G; 147 T; 0 U; 0 Other;
XX      QY      Query Match      46.7%; Score 28; DB 10; Length 717;
XX      Db      Best Local Similarity 66.7%; Pred. No. 6.5;
XX      Matches 40; Conservative 0; Mismatches 20; Indels 0; Gaps 0
XX      1 GCGAGTACTGCACCACCGGACATGGCGTTAACATGACGACTGTATACAGACCTGCACCT 60
XX      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
XX      604 GCGAGTGTAGTACCATCAACACAGCGCTGTACACACACCAACCTGCACCAACAAACACCT 663

```

Search completed: January 15, 2005, 20:36:07
job time : 172.5 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2005, 17:40:52 : Search time 722.8 Seconds
(without alignments)
3925.542 Million cell updates/sec

Title: US-10-057-136-4

Perfect score: 60
Sequence: 1 ggacgactacgcacacacgcgc.....ctgatacaagacctgcacct 60

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 2364849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl: *
1: gb_ba: *
2: gb_hgt: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_seg: *
12: gb_sy: *
13: gb_un: *
14: gb_vi: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	2297	6	BD272907 A recombi
2	36.6	61.0	525	6	BD225141 Medicinal
3	36.6	61.0	891	6	BD225147 Medicinal
4	35	58.3	1371	6	BD225144 Medicinal
5	35	58.3	1737	6	BD225150 Medicinal
6	30	50.0	80009	3	AF226688 Bombyx mo
7	29.2	48.7	132	6	AX180276 Sequence
8	29.2	48.7	144	6	AX180275 Sequence
9	29.2	48.7	557	3	AB017362 Bombyx mo
10	29.2	48.7	2008	3	S74439 silk fibro
11	28.6	47.7	10701	1	AE006311 Lactococc
12	28.6	47.7	227395	10	AC119248 Mus muscu
13	28.6	47.7	256754	2	AC109139 Mus muscu
14	28.6	47.3	69233	9	AC109778 Homo sapi
15	28.4	47.3	96345	9	AC117438 Homo sapi
16	28.4	47.3	147727	2	AC011950 Homo sapi
17	28	46.7	693	8	AK058815 Oryza sat
18	28	46.7	1069	12	AF283517 Synthetic
19	28	46.7	2400	1	CTCIPBA X68233 C. thermocel

20	28	46.7	5573	1	CLOCIPA	L08665 Clostridium
21	28	46.7	179714	8	AP002743	AP002743 Oryza sat
22	27.8	46.3	5220	1	AF479753	AF479753 Lactobaci
23	27.6	46.0	324	3	BMFIBMR	X13869 B. mori MRN
24	27.6	46.0	324	3	BMOFIBA	M35378 B. mori silk
25	27.6	46.0	224322	2	CR450711	CR450711 Dario rer
26	27.6	46.0	234477	2	AC129063	AC129063 Rattus no
27	27.6	46.0	239726	2	AC097728	AC097728 Rattus no
28	27.6	46.0	248552	2	AC095542	AC095542 Rattus no
29	27.4	45.7	984	2	AR375339	AR375339 Sequence
30	27.4	45.7	1073	10	AB032764	AB032764 Mus muscu
31	27.4	45.7	216885	2	CR382328	CR382328 Mus muscu
32	27.2	45.3	67177	2	AC103398	AC103398 Mus muscu
33	27.2	45.0	120	6	AX192396	AX192396 Sequence
34	27	45.0	120	6	BD000571	BD000571 Human pol
35	27	45.0	120	6	BD000571	BD000571 Human pol
36	27	45.0	1455	6	AX959912	AX959912 Sequence
37	27	45.0	1457	6	AX959914	AX959914 Sequence
38	27	45.0	1572	6	AX093798	AX093798 Sequence
39	27	45.0	1774	6	AX959684	AX959684 Sequence
40	27	45.0	1774	6	AX959912	AX959912 Sequence
41	27	45.0	1804	6	AR492318	AR492318 Sequence
42	27	45.0	1804	6	AX335367	AX335367 Sequence
43	27	45.0	1804	9	HUMMUCAB	J05581 Human polym
44	27	45.0	1834	12	AF423030	AF423030 Synthetic
45	27	45.0	1835	6	AX959918	AX959918 Sequence

ALIGNMENTS

RESULT 1	BD272907	2297 bp	DNA	linear	PAT 17-JUL-2003
LOCUS	BD272907	A recombinant vector expressing multiple constitutively molecules and uses thereof.			
DEFINITION	BD272907	A recombinant vector expressing multiple constitutively molecules and uses thereof.			
ACCESSION	BD272907	BD272907.1 GI:33082675			
VERSION	BD272907	JP 2002531133-A/1.			
KEYWORDS	JP 2002531133-A/1.	synthetic construct			
SOURCE	JP 2002531133-A/1.	synthetic construct			
ORGANISM	JP 2002531133-A/1.	artificial sequences.			
REFERENCE	JP 2002531133-A/1.	1 (bases 1 to 2297)			
AUTHORS	Schlom, J., Hodge, J. and Panicali, D.	A recombinant vector expressing multiple constitutively molecules and uses thereof			
TITLE	Patent: JP 2002531133-A 1 24-SEP-2002;	THE UNITED STATES OF AMERICA, THERION BIOLOGICS CORP			
JOURNAL	JP 2002531133-A/1	Artificial Sequence			
COMMENT	JP 2002531133-A/1	Artificial Sequence			
PN	JP 2002531133-A/1	Artificial Sequence			
PD	24-SEP-2002	Artificial Sequence			
PF	12-NOV-1999 JP 2000586927	Artificial Sequence			
PR	09-DEC-1998 US 60/111582	Artificial Sequence			
PI	JEFFREY SCHLOM, JAMES HODGE, DENNIS PANICALI	Artificial Sequence			
PC	C12N15/02, A61K35/12, A61K35/74, A61K35/76, A61K38/00, A61K39/00, A61K39/12, A61K39/21, A61K39/235, A61K39/245, A61K39/275, A61K39/29, A61K48/00.	Artificial Sequence			
PC	A61K39/12, A61K39/21, A61K39/235, A61K39/245, A61K39/275, A61K39/29, A61K48/00.	Artificial Sequence			
PC	A61P1/04, A61P31/04, A61P31/10, A61P31/12, A61P35/00, A61P37/02, PC A61P37/06.	Artificial Sequence			
PC	C12N1/15, C12N1/19, C12N1/21, C12N5/10, C12N7/00, C12Q1/02, G01N3/53, C12N15/00.	Artificial Sequence			
PC	C12N5/00, A61K37/02	Artificial Sequence			
CC	Description of Artificial Sequence: VECTOR SEQUENCE FH	Key			
FT	Location/Qualifiers	Location/Qualifiers			
FT	1. .2297	Location/Qualifiers			
FEATURES	Location/Qualifiers	Location/Qualifiers			
source	1. .2297	Location/Qualifiers			
ORIGIN	/organism="synthetic construct"	/organism="synthetic construct"			
	/mol_type="genomic DNA"	/mol_type="genomic DNA"			
	/db_xref="taxon:32630"	/db_xref="taxon:32630"			

Query Match 100.0%; Score 60; DB 6; Length 2297;
Best Local Similarity 100.0%; Pred. No. 4.3e-09;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCAGTACTGCACACCGGCACATGGCGTAATCATGACCTGATATCAAGACCTGCACCT 60
|||||
Db 406 GGCAGTACTGCACACCGGCACATGGCGTAATCATGACCTGATATCAAGACCTGCACCT 465

RESULT 2
BD225141
LOCUS
DEFINITION
525 bp DNA linear PAT 17-JUL-2003
Medicinal composition having antitumor effect and containing DNA
BD225141
BD225141
VERSION
BD225141.1 GI:33034911
KEYWORDS
JP 2002528519-A/2.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 525)
Pallente,D., Massimo,A.M.D. and Desantis,R.
Medicinal composition having antitumor effect and containing DNA
encoding antigenic protein
Patent: JP 2002528519-A 2 03-SEP-2002;
MENARINI RICERCHE SPA
OS Homo sapiens (human)
PN JP 2002528519-A/2
PD 03-SEP-2002
PF 18-OCT-1999 JP 2000579265
PR 30-OCT-1998 IT MI98A002330
PI DINO PALLEENTE,ANNA MARIA D MASSIMO,RITA DESANTIS PC
A61K38/00,A61K39/76,A61K39/00,A61K48/00,A61P35/00,C12N15/09, PC
A61K37/02,
PC C12N15/00
CC Medicinal composition having antitumor effect and containing
CC DNA encoding
CC antigenic protein
CC key
FH Key
FT source
PT Location/Qualifiers
1..525
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

FEATURES
source
1..525
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 61.0%; Score 36.6; DB 6; Length 525;
Best Local Similarity 76.3%; Pred. No. 0.24;
Matches 45; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 1 GGCAGTACTGCACACCGGCACATGGCGTAATCATGACCTGATATCAAGACCTGCACCT 59
|||||
Db 256 GGAAGTACTGCTCCACACGACGATGTTACTCGGCTCCGATATCAAGACCTGCACCT 314

RESULT 3
BD225147
LOCUS
DEFINITION
891 bp DNA linear PAT 17-JUL-2003
Medicinal composition having antitumor effect and containing DNA
BD225147
BD225147
VERSION
BD225147.1 GI:33034917
KEYWORDS
JP 2002528519-A/8.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 891)
Pallente,D., Massimo,A.M.D. and Desantis,R.

REFERENCE
1 (bases 1 to 891)
Pallente,D., Massimo,A.M.D. and Desantis,R.

AUTHORS
Pallente,D., Massimo,A.M.D. and Desantis,R.

TITLE
Medicinal composition having antitumor effect and containing DNA
encoding antigenic protein

JOURNAL
PATENT: JP 2002528519-A 5 03-SEP-2002;
MENARINI RICERCHE SPA
OS Homo sapiens (human)
PN JP 2002528519-A/5
PD 03-SEP-2002
PF 18-OCT-1999 JP 2000579265
PR 30-OCT-1998 IT MI98A002330
PI DINO PALLEENTE,ANNA MARIA D MASSIMO,RITA DESANTIS PC
A61K38/00,A61K39/76,A61K39/00,A61K48/00,A61P35/00,C12N15/09, PC
A61K37/02,
PC C12N15/00
CC Medicinal composition having antitumor effect and containing
CC DNA encoding
CC antigenic protein
CC key
FH Key
FT source
PT Location/Qualifiers
1..1371
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

TITLE
JOURNAL
COMMENT
OS Homo sapiens (human)
PN JP 2002528519-A/8
PD 03-SEP-2002
PF 18-OCT-1999 JP 2000579265
PR 30-OCT-1998 IT MI98A002330
PI DINO PALLEENTE,ANNA MARIA D MASSIMO,RITA DESANTIS PC
A61K38/00,A61K39/76,A61K39/00,A61K48/00,A61P35/00,C12N15/09, PC
A61K37/02,
PC C12N15/00
CC Medicinal composition having antitumor effect and containing
CC DNA encoding
CC antigenic protein
CC key
FH Key
FT source
PT Location/Qualifiers
1..891
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

FEATURES
source
1..891
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 61.0%; Score 36.6; DB 6; Length 891;
Best Local Similarity 76.3%; Pred. No. 0.23;
Matches 45; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 1 GGCAGTACTGCACACCGGCACATGGCGTAATCATGACCTGATATCAAGACCTGCACCT 59
|||||
Db 622 GGAAGTACTGCTCCACACGACGATGTTACTCGGCTCCGATATCAAGACCTGCACCT 680

RESULT 4
BD225144
LOCUS
DEFINITION
1371 bp DNA linear PAT 17-JUL-2003
Medicinal composition having antitumor effect and containing DNA
BD225144
BD225144
VERSION
BD225144.1 GI:33034914
KEYWORDS
JP 2002528519-A/5.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1371)
Pallente,D., Massimo,A.M.D. and Desantis,R.
Medicinal composition having antitumor effect and containing DNA
encoding antigenic protein
Patent: JP 2002528519-A 5 03-SEP-2002;
MENARINI RICERCHE SPA
OS Homo sapiens (human)
PN JP 2002528519-A/5
PD 03-SEP-2002
PF 18-OCT-1999 JP 2000579265
PR 30-OCT-1998 IT MI98A002330
PI DINO PALLEENTE,ANNA MARIA D MASSIMO,RITA DESANTIS PC
A61K38/00,A61K39/76,A61K39/00,A61K48/00,A61P35/00,C12N15/09, PC
A61K37/02,
PC C12N15/00
CC Medicinal composition having antitumor effect and containing
CC DNA encoding
CC antigenic protein
CC key
FH Key
FT source
PT Location/Qualifiers
1..1371
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

Query Match	58.3%	Score 35	DB 6	Length 1371
Best Local Similarity	74.6%	Pred. No. 0.72		
Matches 44	Conservative 0	Mismatches 15	Indels 0	Gaps 0
QY	1	GGCAGTACTGCACCGCAGCATGCGCTAATCATCAGCAGCCTGATACAGACCTGCACC	59	
DB	325	GGAAGTACCGCTCCACCGACACGCGGTGTACTCGGCTCGGATACAGAGCGCGGCCCC	383	
RESULT 5				
LOCUS	BD225150	1737 bp	DNA	linear
DEFINITION	Medical composition having antitumor effect and containing DNA encoding antigenic protein.			
ACCESSION	BD225150			
VERSION	BD225150.1	GI:33034920		
KEYWORDS	JP 2002528519-A/11.			
SOURCE	Human sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
AUTHORS	1 (bases 1 to 1737)			
TITLE	Palente,D., Massimo,A.M.D. and Desantis,R.			
JOURNAL	Medical composition having antitumor effect and containing DNA encoding antigenic protein			
COMMENT	Patent: JP 2002528519-A 11 03-SEP-2002; MENARINI RICERCHE SPA			
OS	Homo sapiens (human)			
PN	JP 2002528519-A/11			
PD	03-SEP-2002			
PF	18-OCT-1999 JP 2000579265			
PI	30-OCT-1998 IT MI98A002330			
PI	DINO PALLENTE, ANNA MARIA D MASSIMO, RITA DESANTIS PC			
A61K38/00	A61K35/76, A61K39/00, A61K48/00, A61P35/00, C12N15/09, PC			
A61K37/02				
PC	C12N15/00			
CC	Medical composition having antitumor effect and containing DNA encoding			
CC	antigenic protein			
CC	key			
PH	source			
FT	1.1737			
FEATURES	Location/Qualifiers			
source	1.1737			
	/organism="Homo sapiens"			
	/mol_type="genomic DNA"			
	/db_xref="taxon:9606"			
ORIGIN				
Query Match	58.3%	Score 35	DB 6	Length 1737
Best Local Similarity	74.6%	Pred. No. 0.7		
Matches 44	Conservative 0	Mismatches 15	Indels 0	Gaps 0
QY	1	GGCAGTACTGCACCGCAGCATGCGCTAATCATCAGCAGCCTGATACAGACCTGCACC	59	
DB	691	GGAAGTACCGCTCCACCGACACGCGGTGTACTCGGCTCGGATACAGAGCGCGGCCCC	749	
RESULT 6				
LOCUS	AF226688	8009 bp	DNA	linear
DEFINITION	Bombyx mori fibroin heavy chain (fib-H) gene, complete cds.			
ACCESSION	AF226688			
VERSION	AF226688.1	GI:8572060		
KEYWORDS	Bombyx mori (domestic silkworm)			
SOURCE	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Preyigota; Neoptera; Endopterygota; Lepidoptera; Glossata; Dityria;			
ORGANISM	Bombycidae; Bombycidae; Bombyx.			
REFERENCE	1 (bases 1 to 8009)			

[illegible]

gene /gene="fabd"
/note="synonym: L0184"
613. .1541
/gene="fabd"
/note="EVIDENCE BY HOMOLOG BIO07.01 FATTY ACID AND
PHOSPHOLIPID METABOLISM. General. FATTY ACID BIOSYNTHESIS"
/codon start=1
/transl_table=1
/product="malonyl CoA-acyl carrier protein transacylase
(EC 2.3.1.39)"
/protein_id="AAK04871.1"
/db_xref="GI:12723692"
/translation="MTKTAPLFGGQGNKLGMRADLYDOYETWKATPDASQALGYDL
RALIDNEEKLNETKYOPAILFTSVAILRLSENGIKPDLVAGLSLGERISALVAGT
IDROEAKLVAKRGQYMTTEAPFTGSGRVAVMNTDPLLEICQKAEFFKGIVSAPN
YNTPAQIVIGSEVADVAVELKENGVRKLIILKVGSPPTAILKPAESKALIELDK
IDRPPEPLPLISNTSAKMENDEKGLTRQVMEPVRFESVETMQKLGATPFIIEVGP
GRVLSGFIKKIKDKNAEIANVENNLASFALINQ"
1627. .2358
/gene="fabg1"
/note="synonym: L0185"
1627. .2358
/gene="fabg1"
/note="EVIDENCE BY HOMOLOG BIO07.01 FATTY ACID AND
PHOSPHOLIPID METABOLISM. General. FIRST REDUCTION STEP IN
THE FATTY ACID BIOSYNTHESIS PATHWAY"
/codon start=1
/transl_table=1
/product="3-oxoacyl-acyl carrier protein reductase (EC
1.1.1.100)"
/protein_id="AAK04872.1"
/db_xref="GI:12723693"
/translation="MEIKKNVFPVGTSTGIGKATLQPAKAGSNLIINGRAISEEL
LAFTYGVAVGIVSGDLSKSDAKQWMAEPIITLGSYDILVNNAGITTDGILSKXSE
EDESILKINTLTARPMTOVTLKPTIRASGATINSSVGLMNGANVAYASAGL
IGTRKSIAREVAARNRVANAAGFLTESMTVELSDVKDANKGQIPMRFCMPBEIA
TATQPLAEQRYMNGVLLTIDGVS"
2399. .3628
/gene="fabf"
/note="synonym: L0186"
2399. .3628
/gene="fabf"
/note="EVIDENCE BY HOMOLOG BIO07.01 FATTY ACID AND
PHOSPHOLIPID METABOLISM. General. FATTY ACID BIOSYNTHESIS"
/codon start=1
/transl_table=1
/product="3-oxoacyl-acyl carrier protein synthase II (EC
2.3.1.41)"
/protein_id="AAK04873.1"
/db_xref="GI:12723694"
/translation="MTNRVVISGYGVSAIGNSPEEFMNNLKAGTKGIGPIITPDABA
TGISVAEVEKPEPPDKYFOKDDARRMDFTSLAAYVALDMMMSGITEENFPDLCG
IMASGIGLEOSQKNSQATVAKGPRKQVAPLYPLAIDNMTAGNALTLGARGVSRAR
VTAAGANSISAPREIRGADAMINAGARAALCEIGIAPANTLTALTEDEKA
CIPDPKRGFPWBGSGVLIILLESLSHAQARGANILAEIVGNTNDAPFMTPEVGG
AENAILALEEGASPADVDYVNAHGTSTHANEETESKAINHVIDGKAVYSTKLHG
HALGATGAIEAVATVOALINQYAPVNAAGTTEDEGMTINVVLEGKPADIMIALSDOF
GFGGNVAVLAFFKWTGE"
3635. .4102
/gene="accb"
/note="synonym: L0187"
3635. .4102
/gene="accb"
/note="EVIDENCE BY HOMOLOG BIO07.01 FATTY ACID AND
PHOSPHOLIPID METABOLISM. General. FIRST STEP IN LONG-CHAIN
FATTY ACID SYNTHESIS"
/codon start=1
/transl_table=1
/product="biotin carboxyl carrier protein of acetyl-CoA
carboxylase (EC 6.4.1.2)"
/protein_id="AAK04874.1"
/db_xref="GI:12723695"
/translation="WNISBYKDLNMQPFGSTILRFERSWKNADGELSFKNEGQALVGAQ

gene /gene="fabz2"
4229. .4663
/note="synonym: L0188"
4229. .4663
/gene="fabz2"
/note="EVIDENCE BY HOMOLOG BIO07.01 FATTY ACID AND
PHOSPHOLIPID METABOLISM. General. FATTY ACID BIOSYNTHESIS"
/codon start=1
/transl_table=1
/product="3R-hydroxymyristoyl-(3R)-hydroxymyristoyl-acyl
carrier protein dehydratase (EC 4.2.1.1)"
/protein_id="AAK04875.1"
/db_xref="GI:12723696"
/translation="MTEVINIVTEIMEALPHRYPFLVDVRVIDIAEDETAKNVTIN
EEFQGHGPOYVMBGVILMALQAAQVLELSKEPKKGLVFAAGMDNVKKQGVIP
GDOLVIAHKFIKRRGPIAVVVEAATVQKLAAGTLPALGR"
5057. .6424
/gene="accC"
/note="synonym: L0189"
5057. .6424
/gene="accC"
/note="EVIDENCE BY HOMOLOG BIO07.01 FATTY ACID AND
PHOSPHOLIPID METABOLISM. General. FIRST STEP IN LONG-CHAIN
FATTY ACID SYNTHESIS"
/codon start=1
/transl_table=1
/product="biotin carboxylase (EC 6.3.4.14)"
/protein_id="AAK04876.1"
/db_xref="GI:12723697"
/translation="MFKKILIANRGEIAVNIIRARRELGVTAIVASTADRALAHHL
ADEAICVGPASTESYLMNNILISAATAQAATHPQGFLSNSKPARLCEEMNIXF
IGSPAKWMDMGDINARAEMIAGAVVTGSGEBEYTTDEAVKIAEIGYPTILNKS
AGCGKGIKRVNNVEELVPAFEAAQAAQAFGAMFIDRMIFPAHIEVQILIGSQ
GNVHILEKPCDSIQRNNOKVLESPVAIGHTIRNETTAANAARAAOHVSYENGTIEF
LHDESGKFEPMENKTRAVOEHPYTERVGEVDIYKQIRVANGNELSKROEDVTFSGH
AIECRINAKPKFIRPARGKITNPLPSGCVGLRDSAMYSQSYIPPYDSMTAKVI
VHGNRPBALMKQKRALIEFDESVITTVDFQLELSDPHVAVGDDTSLGAVPLFE
YOKSD"
6510. .7376
/gene="accD"
/note="synonym: L0180"
6510. .7376
/gene="accD"
/note="EVIDENCE BY HOMOLOG BIO07.01 FATTY ACID AND
PHOSPHOLIPID METABOLISM. General. FATTY ACID BIOSYNTHESIS"
/codon start=1
/transl_table=1
/product="acetyl-CoA carboxylase carboxyl transferase
subunit beta (EC 6.4.1.2)"
/protein_id="AAK04877.1"
/db_xref="GI:12723698"
/translation="MALFOKKYIKINPNRSIIIEKQAEQEVDELPAKCPACKHTIY
QKDLGNKVCPCNDVIRITAKERLAIIVADKOSFVMEFGIESEKNPLDPGYEKLAA
TKARIGKDEAVITGTAITIGQKTAALIMSTFMAAMGVVGBKILRLREYATTEKLP
IIVFTASGARMOEGIMSMQAKSTAAYKRGHNDALFITVUTDPTTSGVTRASPIL
GDIIIAEPQSLIFAGRRVTEQVTRQTLPPDFQKAEFLINHGVDIVKRTLEKQKLA
LLLELHTEVENV"
7369. .8151
/gene="acca"
/note="synonym: L0190"
7369. .8151
/gene="acca"
/note="EVIDENCE BY HOMOLOG BIO07.01 FATTY ACID AND
PHOSPHOLIPID METABOLISM. General. FIRST STEP IN LONG-CHAIN
FATTY ACID SYNTHESIS"
/codon start=1
/transl_table=1
/product="acetyl-CoA carboxylase carboxyl transferase
subunit alpha (EC 6.4.1.2)"
/protein_id="AAK04878.1"
/db_xref="GI:12723699"

/translaction="MSETIDSLTGEVSKNINADIINEARADRIARISADIIIDKFTDF
 EAHGRDLDDDAIVIGIARFNNOAVTIGIOKGRILANLITNEQSPBNRYRRLA
 LMGKREKRPVITPINTGAPGIAEABERGGEALABRLTMSDLKVIKIPITGEG
 GSGGALALANKVITLERNVSVSPSPFATILKMGSRDEALMKITAPHLEEM
 GIVDAIISEENIKRDLTSGLSLTLETTELAEQRYERPKY"
 8280.9422
 /gene="metB2"

Query Match 47.7%; Score 28.6; DB 1; Length 10701;
 Best Local Similarity 72.5%; Pred. No. 70;
 Matches 37; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Db 10 GCACCAACGCGACATGCGCTAACATCAGACCGTATACAGACCTGACCT 60
 2901 GCACCAACGCGACATGCGCTAACATCAGACCGTATACAGACCTGACCT 2851

RESULT 12 AC119248 227395 bp DNA linear ROD 29-APR-2004
 LOCUS Mus musculus chromosome 8, clone RP24-273H15, complete sequence.
 AC119248
 AC119248.7 GI:46849687
 VERSION HTG.
 KEYWORDS Mus musculus (house mouse)
 SOURCE
 ORGANISM

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 227395)
 Birren, B., Nuebaum, C. and Lander, E.
 Mus musculus chromosome 8, clone RP24-273H15
 Unpublished
 2 (bases 1 to 227395)

Birren, B., Linton, L., Nuebaum, C., Lander, E., All, A., Allen, N.,
 Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,
 Boukhgalter, B., Brown, A., Camarata, J., Campiano, A., Chang, J.,
 Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,
 Cook, A., Cooke, P., DeArliano, K., Dewar, K., Diaz, J.S., Dodge, S.,
 Fato, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
 Ginde, S., Goid, S., Goyette, M., Graham, L., Grand-Pierre, N.,
 Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
 Kamat, A., Karatas, A., Kells, C., Larocque, K., Lamazares, R.,
 Lander, T., Lehoczy, J., Levine, R., Margulis, N., Matthews, C.,
 Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,
 McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Meneus, L.,
 Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,
 Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,
 Oliver, J., Peterson, K., Phunhngang, P., Pierre, N., Pollara, V.,
 Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
 Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupack, R.,
 Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
 Strahan, N., Subramanian, A., Talamas, J., Testfaye, S., Theodore, J.,
 Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
 Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE Direct Submission
 JOURNAL Submitted (25-APR-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 227395)
 Birren, B., Nuebaum, C., Lander, E., Abouneillel, A., Allen, N.,
 Anderson, M., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T.,
 Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y.,
 Collymore, A., Cook, A., Cooke, P., Corum, B., DeArliano, K.,
 Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Fato, S.,
 Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S.,
 Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B.,
 Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
 Kamat, A., Karatas, A., Kells, C., Lander, T., Levine, R.,
 Lindblad-Toh, K., Liu, X., Lui, A., Mabbitt, R., Maclean, C.,
 Macdonald, P., Major, J., Manning, T., Matthews, C., McCarthy, M.,
 Meldrum, J., Meneus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J.,
 Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
 O'Neill, D., Oliver, J., Peterson, K., Phunhngang, P., Pierre, N.,

Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P.,
 Roman, J., Schauer, S., Schupack, R., Seaman, S., Severy, P., Smith, C.,
 Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M.,
 Talamas, J., Testfaye, S., Theodore, J., Topham, K., Travers, M.,
 Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X.,
 Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
 Direct Submission
 JOURNAL Submitted (18-MAR-2004) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 4 (bases 1 to 227395)
 Birren, B., Nuebaum, C., Lander, E., Abouneillel, A., Allen, N.,
 Anderson, M., Anderson, S., Arachchi, H.M., Barna, N., Bastien, V.,
 Bloom, T., Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J.,
 Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B.,
 DeArliano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L.,
 Erickson, J., Fato, S., Ferreira, P., Fitzgerald, M., Gage, D.,
 Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N.,
 Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I.,
 Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Lander, T.,
 Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R.,
 Maclean, C., Macdonald, P., Major, J., Manning, T., Matthews, C.,
 McCarthy, M., Meldrum, J., Meneus, L., Mihova, T., Mienga, V.,
 Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C.,
 O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,
 Phunhngang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C.,
 Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupack, R.,
 Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N.,
 Stojanovic, N., Stubbs, M., Talamas, J., Testfaye, S., Theodore, J.,
 Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
 Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L.,
 Zimmer, A. and Zody, M.

TITLE Direct Submission
 JOURNAL Submitted (29-APR-2004) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Apr 29, 2004 this sequence version replaced gi:45544751.

COMMENT All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@road.mit.edu
 Project Information
 Center project name: 125340
 Center clone name: 273_H_15

FEATURES
 source location/Qualifiers
 1..227395
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="8"
 /map="8"
 /clone="RP24-273H15"
 /clone_11b="RPCT-24 Male Mouse BAC"
 1...6920
 /note="wgs and extension
 clone end:5p6"
 8..91
 /rpt_family="CT-rich"
 complement(1114..1557)
 /rpt_family="RME6B"
 2654..2699
 /rpt_family="(CTTG)n"
 2710..3134
 /rpt_family="WTC"
 3135..5529
 /rpt_family="L1P5"
 3821..3842
 /note="single clone coverage"
 5834..6010
 /rpt_family="B2_Mm2"

```

repeat_region 6032..6296
/rpt_family="LIM2"
repeat_region 6426..6461
/rpt_family="(CA)n"
repeat_region complement(6700..6853)
/rpt_family="RLTR11A"
repeat_region 6903..6979
/rpt_family="LIME"
misc_feature 6921..6924
/note="clone boundary
clone_end:SP6
site:MboI"
repeat_region 7330..7378
/rpt_family="(TTTTTG)n"
repeat_region complement(7381..7935)
/rpt_family="Lx"
repeat_region complement(8566..8781)
/rpt_family="MRT1D"
repeat_region complement(9071..9197)
/rpt_family="ORR1D"
repeat_region 9384..9442
/rpt_family="(TTAA)n"
repeat_region 9415..9518
/rpt_family="(TTATA)n"
repeat_region complement(9586..9807)
/rpt_family="URR1B"
repeat_region 10143..10171
/rpt_family="(A)n"
repeat_region 10201..10335
/rpt_family="B1F"
repeat_region 11649..11826
/rpt_family="MER20"
repeat_region 11897..12045
/rpt_family="(TC)n"
repeat_region 12610..12654
/rpt_family="GA-rich"
repeat_region 14438..14460
/rpt_family="(TTTG)n"
repeat_region 15295..15386
/rpt_family="(TAA)n"
repeat_region 15527..15561
/rpt_family="(TTTTG)n"
repeat_region complement(15755..15820)
/rpt_family="B3"
repeat_region complement(16742..16891)
/rpt_family="B1F"
repeat_region complement(16827..16905)
/rpt_family="B4A"
repeat_region 17235..17269
/rpt_family="(TGC)n"
repeat_region complement(18173..18259)
/rpt_family="13b"
repeat_region 18626..20821
/rpt_family="L1_MM"
repeat_region 22232..22273
/rpt_family="(TAA)n"
repeat_region 22855..22915
/rpt_family="(GGGA)n"
repeat_region complement(24184..24315)
/rpt_family="ID_B1"
repeat_region complement(24808..25550)
/rpt_family="RMR1B"
repeat_region 27215..27345
/rpt_family="B4A"
repeat_region 27346..27366
/rpt_family="(CAGAGA)n"
repeat_region 27367..27491
/rpt_family="B4A"
repeat_region complement(29732..29960)
/rpt_family="LIMB4"
repeat_region complement(29941..30530)
/rpt_family="MER82"
repeat_region 30805..30919

```

Query Match 47.7%; Score 28.6; DB 10; Length 227395;
 Best Local Similarity 67.8%; Pred. No. 49;
 Matches 40; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 1 GCGAGTACGTCACCGGACATGCGTAACTACGACCTGTATCAAGACCTGCACC 59
 |||||
 Db 157274 GCCAGCATCTCCACGACGACGACGTCATCGCATGCGTGTATTAACCTGCC 157332
 |||||

RESULT 13
 AC109139/c 256754 bp DNA linear HTG 13-MAR-2004
 LOCUS Mus musculus chromosome 8 clone RP23-47L13 map 8, WORKING DRAFT
 DEFINITION SEQUENCE, 27 unordered pieces.
 AC109139
 AC109139.4 GI:45430241
 HTG: HTGS PHASE1; HTGS DRAFT; HTGS FULLTOP.
 VERSIONS
 KEYWORDS
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 256754)
 Birren, B., Nusbaum, C. and Lander, E.
 Mus musculus chromosome 8, clone RP23-47L13
 Unpublished
 2 (bases 1 to 256754)
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
 Anderson, S., Barna, N., Baetien, V., Boguslavsky, L., Boukhalter, B.,
 Brown, A., Camarata, J., Campoliano, A., Chang, J., Chazaro, B.,
 Chospel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A.,
 Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S.,
 Ferrelia, P., Fitzhugh, M., Gage, D., Galagan, J., Gaidyna, S.,
 Ginde, S., Gord, S., Goyette, M., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
 Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
 Kamat, A., Karatas, A., Kells, C., Lacroque, K., Lamazares, R.,
 Landers, T., Lehoczy, J., Levine, R., Liu, G., Maclean, C.,
 MacDonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M.,
 McDermid, P., McKernan, K., Meidrim, J., Meneus, L., Mihova, T.,
 McEwan, P., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C.,
 Menga, V., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J.,
 Norman, C. H., O'Connor, T., O'Donnell, P., Pollara, V., Raymond, C.,
 Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C.,
 Retter, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J.,
 Rosetti, M., Roy, A., Santos, R., Schauer, S., Schnupack, R., Seaman, S.,
 Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
 Straus, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J.,
 Topham, K., Travers, M., Travis, N., Trifillo, D., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,
 Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE Direct Submission
 JOURNAL Submitted (03-FEB-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 256754)
 REFERENCES
 AUTHORS

Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,
 Anderson, M., Arachchi, H. M., Barna, N., Baetien, V., Bloom, T.,
 Boguslavsky, L., Boukhalter, B., Camarata, J., Chang, J., Chospel, Y.,
 Collymore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K.,
 Diaz, J. S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S.,
 Ferrelia, P., Fitzgerald, M., Gage, D., Galagan, J., Gaidyna, S.,
 Graham, L., Grand-Pierre, N., Hages, N., Hagopian, D., Hagos, B.,
 Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
 Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,
 Lindblad-Toh, K., Liu, X., Lui, A., Mabbitt, R., Maclean, C.,
 MacDonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,
 Meidrim, J., Meneus, L., Mihova, T., Menga, V., Murphy, T., Naylor, J.,
 Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
 O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,
 Rachupka, A., Ramasamy, V., Raymond, C., Retter, R., Rise, C., Rogov, P.,
 Roman, J., Schauer, S., Schnupack, R., Seaman, S., Severy, P., Smith, C.,
 Spencer, B., Stange-Thomann, N., Stojanovic, N., Stuyvers, M.,
 Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M.,
 Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X.,
 Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL
COMMENT

Direct Submission
Submitted (13-MAR-2004) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 13, 2004 this sequence version replaced gi:21326298.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information
Center project name: L14110
Center clone name: 47_L13

----- Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 246913 bases at least Q40
Consensus quality: 250922 bases at least Q40
Consensus quality: 252476 bases at least Q30
Insert size: 275000; agarose-fp
Insert size: 253647; sum-of-contigs
Quality coverage: 5.9 in Q20 bases; agarose-fp
Quality coverage: 6.4 in Q20 bases; sum-of-contigs

----- NOTE: This is a 'working draft' sequence. It currently
* consists of 27 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1244: contig of 1244 bp in length
* 1245 1344: gap of 100 bp
* 1246 2423: contig of 1079 bp in length
* 2424 2523: gap of 100 bp
* 2524 18381: contig of 15858 bp in length
* 18382 18481: gap of 100 bp
* 18482 20692: contig of 2211 bp in length
* 20693 20792: gap of 100 bp
* 20793 22569: contig of 1777 bp in length
* 22570 22669: gap of 100 bp
* 22670 37683: contig of 15014 bp in length
* 37684 37784: gap of 100 bp
* 37785 62795: contig of 25012 bp in length
* 62796 62895: gap of 100 bp
* 62896 117535: contig of 54640 bp in length
* 117536 117635: gap of 100 bp
* 117636 119189: contig of 1554 bp in length
* 119190 119289: gap of 100 bp
* 119290 121556: contig of 2267 bp in length
* 121557 121657: gap of 100 bp
* 121658 122227: contig of 571 bp in length
* 122228 122327: gap of 100 bp
* 122328 123313: contig of 986 bp in length
* 123314 123413: gap of 100 bp
* 123414 128705: contig of 5292 bp in length
* 128706 128805: gap of 100 bp
* 128806 130201: contig of 1396 bp in length
* 130202 130301: gap of 100 bp
* 130302 134936: contig of 4635 bp in length
* 134937 135036: gap of 100 bp
* 135037 138419: contig of 3383 bp in length
* 138420 138519: gap of 100 bp
* 138520 145823: contig of 7304 bp in length
* 145824 145923: gap of 100 bp
* 145924 180246: contig of 34323 bp in length
* 180247 180347: gap of 100 bp
* 180348 183714: contig of 3368 bp in length
* 183715 183814: gap of 100 bp

183815 209464: contig of 25650 bp in length
* 209465 209564: gap of 100 bp
* 209565 214498: contig of 4934 bp in length
* 214499 214598: gap of 100 bp
* 214599 216080: contig of 1482 bp in length
* 216081 216180: gap of 100 bp
* 216181 249862: contig of 3362 bp in length
* 249863 249962: gap of 100 bp
* 249963 253576: contig of 3614 bp in length
* 253577 253676: gap of 100 bp
* 253677 254277: contig of 601 bp in length
* 254278 254377: gap of 100 bp
* 254378 256125: contig of 1748 bp in length
* 256126 256225: gap of 100 bp
* 256226 256754: contig of 529 bp in length.

FEATURES
source
1..256754
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="8"
/map="8"
/clone="RP23-47L13"
/clone_lib="RPC1-23 Female Mouse BAC"
1..1244
/note="assembly_fragment"
misc_feature
1345..2423
/note="assembly_fragment"
misc_feature
2524..18381
/note="assembly_fragment"
misc_feature
18482..20692
/note="assembly_fragment"
misc_feature
20793..22569
/note="assembly_fragment"
misc_feature
22670..37683
/note="assembly_fragment"
misc_feature
37784..62795
/note="assembly_fragment"
misc_feature
62896..117535
/note="assembly_fragment"
misc_feature
117636..119189
/note="assembly_fragment"
misc_feature
119290..121556
/note="assembly_fragment"
misc_feature
121657..122227
/note="assembly_fragment"
misc_feature
122328..123313
/note="assembly_fragment"
misc_feature
123414..128705
/note="assembly_fragment"
misc_feature
128806..130201
/note="assembly_fragment"
misc_feature
130302..134936
/note="assembly_fragment"
misc_feature
135037..138419
/note="assembly_fragment"
misc_feature
138520..145823
/note="assembly_fragment"
misc_feature
145924..180246
/note="assembly_fragment"
misc_feature
180347..183714
/note="assembly_fragment"
misc_feature
183815..209464
/note="assembly_fragment"
misc_feature
209565..214498
/note="assembly_fragment"

Query Match 47.7%; Score 28.6; DB 2; Length 256754;
Best Local Similarity 67.8%; Pred. No. 48;
Matches 40; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
1 GGCAGTACTGCACCGGACATGCGTACATGACGACCTGATACAGACCTGCACC 59

Db 108445 GCCAGCACTCCAGCAGCAGCACTCCATGCGATCAGCAGTGGCTATTAATACCTTGGCC 108387

RESULT 14
AC109778
LOCUS
DEFINITION AC109778 69333 bp DNA linear PRI 31-MAY-2002
Homo sapiens 3 BAC RP11-17013 (Roswell Park Cancer Institute Human
BAC library) complete sequence.
AC109778
AC109778.8 GI:21281345
HTG.
Homo sapiens (human)
SOURCE
ORGANISM
REFERENCE
AUTHORS

1 (bases 1 to 69333)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooke,S.L., Amaralunga,H.C., Are,J.R., Ayele,M., Banks,T.,
Barbata,J., Benton,J., Bimaga,K., Blankenburg,K., Bonini,D.,
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Devila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
DeLaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.U.,
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabib,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,M., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., He,X.,
Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,
Hollins,B., Honsi,F., Howard,S., Huber,J., Hulys,S., Hume,J.,
Jackson,L.E., Jacobson,B., Jia,T., Johnson,R., Jolivet,S.,
Koudan,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Lichteberg,O., Lieu,C., Liu,Y., Liu,W.,
Loulesged,H., Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Mei,J., Meshawari,M., Mapa,P., Martin,R., Martindale,A.,
Martinez,E., Massey,E., Mawhinney,E., McLeod,M.P., Meador,M.,
Mei,G., Metzger,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
Moore,S., Morgan,M., Moorish,T., Morris,S., Moser,M., Neal,D.,
Nelson,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N.,
Nickerson,E., Nwokweto,S., Oguh,M., Okwomun,G., Oragunye,N.,
Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L.,
Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M.,
Rojas,A., Rojupokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S.,
Scott,G., Shen,H., Shoonitari,N., Sisson,I., Sodergren,E.,
Sonake,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Swatek,A.,
Taber,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C.,
Taylor,T., Telford,B., Thomas,N., Thomas,S., Umanu,K., Vasquez,L.,
Vera,V., Villalón,D., Vinson,R., Wang,Q., Williams,G.,
Williamson,A., Wleczek,R., Wooden,S., Worley,K., Wu,C., Wu,Y.,
Wu,Y.F., Zhou,J., Zorrilla,S., Zorrilla,S.L., Weinstein,G. and
Gibbs,R.

Direct Submission
Unpublished
2 (bases 1 to 69333)
Worley,K.C.

Direct Submission
Submitted (07-FEB-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 69333)
Worley,K.C.

Direct Submission
Submitted (26-MAY-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
4 (bases 1 to 69333)
Worley,K.C.

TITLE
JOURNAL
COMMENT

Direct Submission
Submitted (31-MAY-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On May 31, 2002 this sequence version replaced gi:20514664.
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email
gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the
entire insert of this clone. Overlapping regions of clones are only
sequenced and submitted once, so the sequence for the remainder of
the insert may be found in the record for the adjacent clones.
Overlapping clones are noted at the beginning and end of the
Features listing.

ANNOTATION OF FEATURES:
STSs are identified using ePCR (Genome Res. 7:541-550) searches
of a local database that includes entries from dbSTS, GDB, and
local mapping efforts.
Repeats are identified using RepeatMasker (A. Smit and P. Green,
unpublished) for human and mouse sequences.
Genes and Region of sequence similarity are identified by BLAST
(Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the
EST and cDNA sequences. Genes demonstrate at least two exons
flanked by consensus splice sites that maintained sequence
continuity across the splice junctions. Sequences that are not
identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum
standard of double strand coverage with a minimum of 2 clones and 2
clones with no ambiguities or 2 chemistries with a minimum of 2
clones and 3 reads with no ambiguities. If the sequence quality for
a region does not meet this standard, it will be indicated in the
annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality
standards - estimated quality individual bases and measures of base
quality are listed below. Description of the metrics can be found
at URL:
http://gc.bcm.tmc.edu:8088/quality_info/genbank_annotation.html.

QUALSTAT-REPORT.

FEATURES	source
Location/Qualifiers	1. 69333
/organism="Homo sapiens"	
/mol_type="genomic DNA"	
/db_xref="taxon:9606"	
/chromosome="3"	
/clone="RP11-17013"	
/complement(1..2006)	
/note="overlaps bases 1..2006 of clone AC069529"	
repeat_region	1844..1864
/rpt_family="AT rich"	
/complement(2195..2250)	
/rpt_family="MER102"	
repeat_region	2862..2867
/rpt_family="(CA)n"	
repeat_region	2868..2900
/rpt_family="(CA)n"	
repeat_region	3085..3192
/rpt_family="HAL1"	
repeat_region	3193..3504
/rpt_family="AluY"	
repeat_region	3505..3771
/rpt_family="HAL1"	
repeat_region	4324..4496
/rpt_family="HAL1"	
repeat_region	5007..5167
/rpt_family="MER5B"	
repeat_region	complement(5194..5964)
/rpt_family="MER9"	

```

repeat_region      6394. .6704      /rpt_family="AluSx"
                    complement(719. .7304)
repeat_region      /rpt_family="MIR"
                    7410. .7453
repeat_region      /rpt_family="(CA)n"
                    complement(8637. .8956)
repeat_region      /rpt_family="AluSq"
                    complement(9936. .10068)
repeat_region      /rpt_family="L2"
                    10183. .10369
repeat_region      /rpt_family="MIR"
                    13768. .13864
repeat_region      /rpt_family="MIR"
                    complement(13865. .14684)
repeat_region      /rpt_family="L1MA4"
                    14685. .14718
repeat_region      /rpt_family="MIR"
                    14711. .14798
repeat_region      /rpt_family="MIR"
                    15323. .15598
misc_feature
                    /note="unresolved tandem approximately 270 bp sized by PCR
                    products and restriction digest HindIII."
                    /function="unresolved tandem repeat"
repeat_region      15324. .15579
                    /rpt_family="(TA)n"
repeat_region      complement(17947. .18003)
                    /rpt_family="L1MA4"
repeat_region      18006. .18083
                    /rpt_family="L1PA16"
repeat_region      18289. .18362
                    /rpt_family="AluSx"
repeat_region      18723. .18770
                    /rpt_family="AluSq"
repeat_region      20231. .20328
                    /rpt_family="AT_rich"
repeat_region      /rpt_family="MIR"
                    21081. .21108
repeat_region      /rpt_family="AT_rich"
                    22465. .22746
repeat_region      /rpt_family="L1PBa"
                    22771. .22917
repeat_region      /rpt_family="AluDb"
                    22923. .25720
repeat_region      /rpt_family="L1PB1"
                    25727. .25785
repeat_region      /rpt_family="(TA)n"
                    25786. .26092
repeat_region      /rpt_family="L1PB1"
                    26093. .26130
repeat_region      /rpt_family="AT_rich"
                    complement(26136. .26211)
repeat_region      /rpt_family="FAM_C"
                    complement(26212. .26625)
repeat_region      /rpt_family="L1"
                    complement(26607. .27510)
repeat_region      /rpt_family="L1M2"
                    28311. .28492
repeat_region      /rpt_family="L1MB4"
                    28356. .28711

```

Query Match	47.3%	Score 28.4	DB 9	Length 69333
Best Local Similarity	70.4%	Pred. NMo.65		
Matches	38	Conservative	0	Matches 16, Indels 0, Gaps 0
Qy	1	GGCAGTACTGCACCCGCGACATGAGCCTGATACAGACCT	54	
Db	69086	GACATTCTGCACCACTGGCTCATGATGCACTTCGACGACATTAACAAACAT	69139	
RESULT 15				
LOCUS	AC117438	96545 bp	DNA	linear
DEFINITION	Homo sapiens 3.BAC RP11-190F5 (Roswell Park Cancer Institute Human			

ACCESSION	BAC Library) complete sequence.
VERSION	AC117438
KEYWORDS	AC117438.4 GI:21591816
SOURCE	HTG.
ORGANISM	Homo sapiens (human)
REFERENCE	Homo sapiens
AUTHORS	Eukaryotes; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 96545) Muzny,D.N., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,

TITLE	Direct Submission
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 96545)
AUTHORS	Worley, K.C.
TITLE	Direct Submission
JOURNAL	Submitted (10-APR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE	3 (bases 1 to 96545)
AUTHORS	Worley, K.C.
TITLE	Direct Submission
JOURNAL	Submitted (23-JUN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE	4 (bases 1 to 96545)
AUTHORS	Worley, K.C.
TITLE	Direct Submission
JOURNAL	Submitted (26-JUN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT	On Jun 26, 2002 this sequence version replaced gi:21490104. INFORMATION: http://www.hgsc.bcm.tmc.edu/ or email gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts. Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences. Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://www.hgsc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

ml.

FEATURES

Source

```

Location/Qualifiers
1..96545
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="3"
/clone="RP11-190F5"
repeat_region
1380..1401
/rpt_family="(T)n"
complement(2723..3026)
repeat_region
3376..3629
/rpt_family="AluYo"
repeat_region
5318..5564
/rpt_family="MIR"
repeat_region
5754..6117
/rpt_family="L2"
repeat_region
8811..9638
/rpt_family="MLT1B"
repeat_region
10070..10296
/rpt_family="L2"
repeat_region
11148..11297
/rpt_family="MER3"
complement(11148..11297)
repeat_region
12694..13521
/rpt_family="MER5A"
repeat_region
13697..13811
/rpt_family="L2"
repeat_region
13957..14361
/rpt_family="MLT1D"
repeat_region
14414..14469
/rpt_family="L2"
repeat_region
14850..15213
/rpt_family="MLT1E"
complement(15861..15933)
repeat_region
16333..16368
/rpt_family="MIR"
repeat_region
16648..16706

```

```

repeat_region
16823..16845
/rpt_family="MIR"
repeat_region
16845..16899
/rpt_family="(TC)n"
repeat_region
16845..16899
/rpt_family="(TA)n"
repeat_region
16845..16899
/rpt_family="(TA)n"
complement(17142..17251)
/rpt_family="L2"
repeat_region
17832..17893
/rpt_family="L1PA12"
repeat_region
17894..18012
/rpt_family="L1PA13"
repeat_region
18420..18702
/rpt_family="AluSp"
repeat_region
18703..18730
/rpt_family="(CA)n"
repeat_region
19326..19355
/rpt_family="AT-rich"
repeat_region
19364..19413
/rpt_family="MIR"
repeat_region
20916..21157
/rpt_family="AluSg"
complement(21560..21933)
/rpt_family="MSTC"
repeat_region
24710..25383
/rpt_family="L2"
complement(25880..26180)
/rpt_family="AluYb8"
complement(26723..26814)
/rpt_family="MER81"
complement(27544..27741)
/rpt_family="MER8"
repeat_region
28231..28707
/rpt_family="MLT1D"
repeat_region
28957..29110
/rpt_family="L1ME"
repeat_region
29557..29728
/rpt_family="AluYo/FRAM"
repeat_region
29780..29814
/rpt_family="(TTCA)n"
repeat_region
29965..30092
/rpt_family="MIR"
repeat_region
31846..31876
/rpt_family="AT-rich"
repeat_region
32282..32318
/rpt_family="(TG)n"
repeat_region
33871..33910
/rpt_family="(TTTAA)n"
complement(33913..34060)
/rpt_family="L1MA4"
repeat_region
34065..34144
/rpt_family="L1MA4"
repeat_region
34145..34345

```

Query Match 47.3%; Score 28.4; DB 9; Length 96545;
 Best Local Similarity 70.4%; Pred. No. 63; 16; Indels 0; Gaps 0;
 Matches 38; Conservative 0; Mismatches

QY 1 GGCAGTACTGCACCGGCACATGCGCTAATCAGCACTTGATCAAGACT 54
 DB 94789 GACATTCCTGCACACTGCGCTATGATGCACTTCAAGCAATTAACAACAT 94736

Search completed: January 15, 2005, 22:36:45
 Job time : 725.8 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2005, 17:40:52 : Search time 722.8 Seconds
(without alignment)
3925.542 Million cell updates/sec

Title: US-10-057-136-2

Perfect score: 60
Sequence: 1 GGCCTCACCGCCCCCGAGCAGGTCGACCTCGGCGGAGACACGAGCGGCGCCG 60

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 10%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_hgt:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_rc:*
11: gb_srs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	60	100.0	120	6	AX192396 Sequence
2	60	100.0	120	6	BD000571 Human pol
3	60	100.0	1414	12	AF423031 Synthetic
4	60	100.0	1455	6	CQ715242 Sequence
5	60	100.0	1457	6	AX959914 Sequence
6	60	100.0	1774	6	AX959684 Sequence
7	60	100.0	1774	6	AX959912 Sequence
8	60	100.0	1804	6	AR492318 Sequence
9	60	100.0	1804	6	AX335367 Sequence
10	60	100.0	1804	6	AX335367 Sequence
11	60	100.0	1834	12	HUMMUCAB AF423030
12	60	100.0	1835	6	AX959918 Sequence
13	60	100.0	2135	6	AX959916 Sequence
14	60	100.0	2297	6	BD272907 A recombi
15	60	100.0	4139	6	CQ834017 Sequence
16	60	100.0	4139	6	AX334899 Sequence
17	60	100.0	4139	6	AX335372 Sequence
18	60	100.0	4139	6	AX336712 Sequence
19	60	100.0	4139	6	AX409474 Sequence

20	60	100.0	4139	6	AX440481 Sequence
21	60	100.0	4139	9	J05582 Human pancr
22	60	100.0	7188	9	AY463543 Homo sapi
23	60	100.0	8181	6	AX406624 Sequence
24	60	100.0	8181	6	W61170 Human polyM
25	60	100.0	8186	6	AR492320 Sequence
26	60	100.0	133525	6	AL713999 Human DNA
27	58.4	97.3	120	6	AX192397 Sequence
28	58.4	97.3	120	6	BD000572 Human pol
29	58.4	97.3	572	6	AR492319 Sequence
30	58.4	97.3	572	6	AR492319 Sequence
31	58.4	97.3	1721	6	CO771290 Human Breas
32	58.4	97.3	1721	6	AR492306 Sequence
33	58.4	97.3	1721	6	AX335860 Sequence
34	58.4	97.3	1721	6	AX440427 Sequence
35	58.4	97.3	1721	6	AX587588 Sequence
36	58.4	97.3	1721	6	AX587588 Sequence
37	58.4	97.3	1971	6	AX963157 Human mRNA
38	58.4	97.3	2037	6	AX963157 Sequence
39	58.4	97.3	2238	9	HSSEBTA X52228 Human mRNA
40	57	95.0	491	9	HUMEPISIBI M32739 Human epist
41	57	95.0	518	6	AR492402 Sequence
42	57	95.0	518	9	HUMEPISIAL M32738 Human epist
43	56.8	94.7	180	6	E08763 CDNA encodi
44	56.8	94.7	1572	6	AX093798 Sequence
45	55.4	92.3	1320	6	E08764 CDNA encodi

ALIGNMENTS

RESULT 1	AX192396/c	Sequence 2 from Patent EP1103623.	120 bp	DNA	linear	PAT 15-AUG-2001
LOCUS	AX192396					
DEFINITION	Sequence 2 from Patent EP1103623.					
ACCESSION	AX192396					
VERSION	AX192396.1	GI:15210363				
KEYWORDS						
SOURCE	Homo sapiens (human)					
ORGANISM	Homo sapiens					
REFERENCE	1 Taylor-Papadimitriou,J., Burchell,J. and Gendler,S. Human mucin core protein: nucleic acid probes, peptide fragments and antibodies thereto, and uses thereof in diagnostic and therapeutic methods Patent: EP 1103623-A 2 30-MAY-2001;					
JOURNAL	IMPERIAL CANCER RESEARCH TECHNOLOGY LIMITED (GB)					
FEATURES	source	1..120				
ORIGIN	1..120	/organism="Homo sapiens"				
		/mol_type="unassigned DNA"				
		/db_xref="taxon:9606"				
Query Match	100.0%; Score 60; DB 6; Length 120;					
Best Local Similarity	100.0%; Pred. No. 0.0025;					
Matches	60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
CY	1 GGCCTCACCGCCCCCGAGCAGGTCGACCTCGGCGGAGACACGAGCGGCGCCG 60					
DB	87 GGCCTCACCGCCCCCGAGCAGGTCGACCTCGGCGGAGACACGAGCGGCGCCG 28					
RESULT 2	BD000571/c	Human polyMorphic epithelial mucin core protein and nucleic acid encoding the protein.	120 bp	DNA	linear	PAT 31-JAN-2002
LOCUS	BD000571					
DEFINITION	Human polyMorphic epithelial mucin core protein and nucleic acid encoding the protein.					
ACCESSION	BD000571					
VERSION	BD000571.1	GI:18623684				
KEYWORDS	JP 2000333675-A/2.					

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
AUTHORS 1 (bases 1 to 120)
TITLE Papadimitrov, J.T., Jendora, S. and Bachieri, J.
JOURNAL Human polymorphic epithelial mucin core protein and nucleic acid Patent: JP 2000333675-A 2 05-DEC-2000;
COMMENT IMPERIAL CANCER RESEARCH TECHNOLOGY LTD
OS Homo sapiens (human)
PN JP 2000333675-A/2
PD 05-DEC-2000
PF 26-APR-2000 JP 2000125724
PR 07-JAN-1987 GB 8700279 07-JAN-1987 GB 8700279 PR
22-APR-1987 US 041306, 03-NOV-1987 GB 8726172 PI JOYCE
TAYLOR PAPADIMITROV, SANDRA JENDORA, JOY BACHIERU PC
C12N15/02, A61K38/00, A61K39/395, A61K49/00, A61P35/00, PC
C07K14/47,
PC C07K16/44, C12N5/10, C12P21/08 // (C12N15/02, C12R1:91), (C12N5/10, C07R1:91),
PC C12R1:91),
PC C12N5/00, A61K37/02, C12N5/00, (C12N15/00, C12R1:91), (C12N5/00, C12R1:91)
CC
FH
FT
Key Location/Qualifiers
Location/Qualifiers
1..120
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

Query Match 100.0%; Score 60; DB 6; Length 120;
Best Local Similarity 100.0%; Pred. No. 0.0025;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 87 GGCTCCACCGCCCGCCGACGAGTGTACCTCGGCGCCGAGACCAAGCGGCGCCG 60
GGCTCCACCGCCCGCCGACGAGTGTACCTCGGCGCCGAGACCAAGCGGCGCCG 28

RESULT 3
AF423031 1414 bp mRNA linear SYN 10-JUL-2003
LOCUS
DEFINITION Synthetic construct Homo sapiens mucin variant MUC1-CT58 (MUC1)
ACCESSION AF423031
VERSION AF423031.1 GI:19338621
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source

synthetic construct
synthetic construct
artificial sequences.
1 (bases 1 to 1414)
Hinojosa-Kurtzberg, A.M., Johansson, M.E., Madsen, C.S., Hansson, G.C. and Gendler, S.J.
Novel MUC1 splice variants contribute to mucin overexpression in CPTN-deficient mice
Am. J. Physiol. Gastrointest. Liver Physiol. 284 (5), G853-G862 (2003)
22570517
12529261
2 (bases 1 to 1414)
Hinojosa-Kurtzberg, A.M. and Gendler, S.J.
Direct Submission
Submitted (21-SEP-2001) Biochemistry, Mayo Clinic, 13400 E. Shea Boulevard, Scottsdale, AZ 85259, USA
Location/Qualifiers
1..1414
/organism="synthetic construct"
/mol_type="mRNA"
/db_xref="taxon:32630"

gene
CDS
/note="Homo sapiens gene in transgenic Mus musculus C57BL/6; isolated from intestinal mucosa"
1..1414
/gene="MUC1"
1..1386
/gene="MUC1"
/note="alternatively spliced; contains exon 6b resulting in variant carboxy-terminal domain; lacks sites for beta-catenin and Grb2 interactions; derived from Homo sapiens"
/codon_start=1
/transl_table=1
/product="mucin variant MUC1-CT58"
/protein_id="AAL86735.1"
/db_xref="GI:19338622"
/translation="MTPTGSPFFLLLLTVLVVVGSGHASTRGGEKTSATGRSSVPSSTKNAVMTSSVLSHSGSGSSTGOGDTLAPATPAGSAATGCDYTSPTVRPALGSTPPPAHDVTSAPDNKRAPGSTAPAGVTSAPPTTPAGSTAPAGVTSAPDNRPALGSTAPPHVNTVTSAGSASGASTLVNNGTSARATVTPASKSTPFSIPISSHSDPTTLASHSTKTDASTHSTVTPPLTSSNHSSTPQSTGVSPFLSFHISLQFNSSLDEPSTDYVGELEORDISEMFLQIKQGGFLGSLNIFRSGSVVQLTLAIREGTINVDHVEFQNYKTEAASRYNLTISDVSVDVFPFSAQSGAGVPMGIALVLVCLVA LAIYLIALAVCCCRKNVGGDLIFPARDTYHPMSXPYTHTRGYVPSPTDSRPIE KEMVRDRERLA"

ORIGIN

Query Match 100.0%; Score 60; DB 12; Length 1414;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 385 GGCTCCACCGCCCGCCGACGAGTGTACCTCGGCGCCGAGACCAAGCGGCGCCG 444
GGCTCCACCGCCCGCCGACGAGTGTACCTCGGCGCCGAGACCAAGCGGCGCCG 444

RESULT 4
CQ715242 1455 bp DNA linear PAT 03-FEB-2004
LOCUS
DEFINITION Sequence 1176 from Patent WO02068579.
ACCESSION CQ715242
VERSION CQ715242.1 GI:42276099
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source

Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
Kits, such as nucleic acid arrays, comprising a majority of humanexons or transcripts, for detecting expression and other uses thereof
Patent: WO 02068579-A 1176 06-SEP-2002;
PE Corporation Location (NY) (US)
Location/Qualifiers
1..1455
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 60; DB 6; Length 1455;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 412 GGCTCCACCGCCCGCCGACGAGTGTACCTCGGCGCCGAGACCAAGCGGCGCCG 471
GGCTCCACCGCCCGCCGACGAGTGTACCTCGGCGCCGAGACCAAGCGGCGCCG 471

RESULT 5
AX959914 1457 bp DNA linear PAT 14-JAN-2004
LOCUS
DEFINITION Sequence 19 from Patent WO03100060.

ACCESSION AX959914
VERSION AX959914.1 GI:40880143
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.
REFERENCE 1
AUTHORS Burden,N.G., Ellis,J.H. and Hamblin,P.A.
TITLE Muc-1 antigen with reduced number of vntz repeat units
JOURNAL Patent: WO 03100060-A 19 04-DEC-2003;
GLAXO GROUP LIMITED (GB)
FEATURES
source 1.1457
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 60; DB 6; Length 1457;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCTCCACCGCCCCCGCCAGCCCGGTGTCACTCGGCCCCGAGACACAGCGCCCGG 60
|||||
403 GGCTCCACCGCCCCCGCCAGCCCGGTGTCACTCGGCCCCGAGACACAGCGCCCGG 462
RESULT 6
AX959684 1774 bp DNA linear PAT 14-JAN-2004
LOCUS
DEFINITION Sequence 10 from Patent WO03099193.
ACCESSION AX959684
VERSION AX959684.1 GI:40880030
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.
REFERENCE 1
AUTHORS Burden,N. and Hamblin,P.
TITLE Vaccines
JOURNAL Patent: WO 03099193-A 10 04-DEC-2003;
GLAXO GROUP LIMITED (GB)
FEATURES
source 1.1774
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 60; DB 6; Length 1774;
Best Local Similarity 100.0%; Pred. No. 0.001;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCTCCACCGCCCCCGCCAGCCCGGTGTCACTCGGCCCCGAGACACAGCGCCCGG 60
|||||
462 GGCTCCACCGCCCCCGCCAGCCCGGTGTCACTCGGCCCCGAGACACAGCGCCCGG 521
RESULT 7
AX959912 1774 bp DNA linear PAT 14-JAN-2004
LOCUS
DEFINITION Sequence 17 from Patent WO03100060.
ACCESSION AX959912
VERSION AX959912.1 GI:40880142
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.
REFERENCE 1
AUTHORS Burden,N.G., Ellis,J.H. and Hamblin,P.A.

TITLE Muc-1 antigen with reduced number of vntz repeat units
JOURNAL Patent: WO 03100060-A 17 04-DEC-2003;
GLAXO GROUP LIMITED (GB)
FEATURES
source 1.1774
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 60; DB 6; Length 1774;
Best Local Similarity 100.0%; Pred. No. 0.001;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCTCCACCGCCCCCGCCAGCCCGGTGTCACTCGGCCCCGAGACACAGCGCCCGG 60
|||||
462 GGCTCCACCGCCCCCGCCAGCCCGGTGTCACTCGGCCCCGAGACACAGCGCCCGG 521
RESULT 8
AR492318 1804 bp DNA linear PAT 15-MAY-2004
LOCUS
DEFINITION Sequence 17 from patent US 6716627.
ACCESSION AR492318
VERSION AR492318.1 GI:47260892
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1804)
AUTHORS Dobie,K.W.
TITLE Antisense modulation of mucin 1, transmembrane expression
JOURNAL Patent: US 6716627-A 17 06-APR-2004;
FEATURES
source 1.1804
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 100.0%; Score 60; DB 6; Length 1804;
Best Local Similarity 100.0%; Pred. No. 0.001;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCTCCACCGCCCCCGCCAGCCCGGTGTCACTCGGCCCCGAGACACAGCGCCCGG 60
|||||
457 GGCTCCACCGCCCCCGCCAGCCCGGTGTCACTCGGCCCCGAGACACAGCGCCCGG 516
RESULT 9
AX335367 1804 bp DNA linear PAT 09-JAN-2002
LOCUS
DEFINITION Sequence 5876 from Patent WO0194629.
ACCESSION AX335367
VERSION AX335367.1 GI:18126086
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.
REFERENCE 1
AUTHORS Young,P.E., Augustus,M., Carter,K.C., Edner,R., Endress,G.,
Horrigan,S., Soper,D.R. and Weaver,Z.
TITLE Cancer gene determination and therapeutic screening using signature
JOURNAL Patent: WO 0194629-A 5876 13-DEC-2001;
Avalon Pharmaceuticals (US)
FEATURES
source 1.1804
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN

Query Match	100.0%;	Score 60;	DB 9;	Length 1804
Best Local Similarity	100.0%;	Pred. No.	0.001;	

RESULT 12

AX959918
LOCUS AX959918 1835 bp DNA linear PAT 14-JAN-2004
DEFINITION Sequence 23 from Patent WO03100060.
ACCESSION AX959918
VERSION AX959918.1 GI:40880145
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 Burden,N.G., Ellis,J.H. and Hamblin,P.A.
Muc-1 antigen with reduced number of vntx repeat units
Patent: WO 03100060-A 23 04-DEC-2003;
JOURNAL GLAXO GROUP LIMITED (GB)
FEATURES
source 1. 1835
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 60; DB 6; Length 1835;
Best Local Similarity 100.0%; Pred. No. 0.00099;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCTCACCAGCCCGCCAGCGGTGTCACTCGCGCCGACACACGCGCGCCCG 60
|||||
462 GGCTCACCAGCCCGCCAGCGGTGTCACTCGCGCCGACACACGCGCGCCCG 521
|||||
RESULT 13
AX959916
LOCUS AX959916 2135 bp DNA linear PAT 14-JAN-2004
DEFINITION Sequence 21 from Patent WO03100060.
ACCESSION AX959916
VERSION AX959916.1 GI:40880144
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 Burden,N.G., Ellis,J.H. and Hamblin,P.A.
Muc-1 antigen with reduced number of vntx repeat units
Patent: WO 03100060-A 21 04-DEC-2003;
JOURNAL GLAXO GROUP LIMITED (GB)
FEATURES
source 1. 2135
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 60; DB 6; Length 2135;
Best Local Similarity 100.0%; Pred. No. 0.00094;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCTCACCAGCCCGCCAGCGGTGTCACTCGCGCCGACACACGCGCGCCCG 60
|||||
462 GGCTCACCAGCCCGCCAGCGGTGTCACTCGCGCCGACACACGCGCGCCCG 521
|||||
RESULT 14
BD272907 2297 bp DNA linear PAT 17-JUL-2003
LOCUS BD272907
DEFINITION A recombinant vector expressing multiple constimulatory molecules
and uses thereof.
ACCESSION BD272907
VERSION BD272907.1 GI:33082675
KEYWORDS JP 2002531133-A/1.
SOURCE synthetic construct
ORGANISM synthetic construct

Artificial sequences.
1 (bases 1 to 2297)
Schlom,J., Hodge,J. and Panticali,D.
A recombinant vector expressing multiple constimulatory molecules
and uses thereof
Patent: JP 2002531133-A 1 24-SEP-2002;
JOURNAL THE UNITED STATES OF AMERICA, THERION BIOLOGICS CORP
OS Artificial Sequence
COMMENT PN JP 2002531133-A/1
PD 24-SEP-2002
PF 12-NOV-1999 JP 2000586927
PR 09-DEC-1998 US 60/111582
PI JEFFREY SCHLOM, JAMES HODGE, DENNIS PANICALI
PC C12N15/02,A61K35/12,A61K35/74,A61K35/76,A61K39/00,
PC A61K39/12,A61K39/21,A61K39/235,A61K39/245,A61K39/275,A61K39/
PC 29,A61K48/00.
PC A61P1/04,A61P31/04,A61P31/10,A61P31/12,A61P35/00,A61P37/02, PC
A61P37/06,
PC C12N1/15,C12N1/19,C12N1/21,C12N5/10,C12N7/00,C12Q1/02,G01N33/
PC 53,C12N15/00,
PC C12N5/00,A61K37/02
CC Description of Artificial Sequence: VECTOR SEQUENCE FH Key
Location/Qualifiers
FT source 1. 2297
/organism="Artificial Sequence".
FEATURES
source 1. 2297
Location/Qualifiers
/organism="Artificial Sequence"
/mol_type="synthetic construct"
/db_xref="taxon:32630"
ORIGIN
Query Match 100.0%; Score 60; DB 6; Length 2297;
Best Local Similarity 100.0%; Pred. No. 0.00092;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCTCACCAGCCCGCCAGCGGTGTCACTCGCGCCGACACACGCGCGCCCG 60
|||||
346 GGCTCACCAGCCCGCCAGCGGTGTCACTCGCGCCGACACACGCGCGCCCG 405
|||||
RESULT 15
CQ834017 4139 bp DNA linear PAT 29-JUL-2004
LOCUS CQ834017
DEFINITION Sequence 53 from Patent EP1439393.
ACCESSION CQ834017
VERSION CQ834017.1 GI:50833622
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 Asle,J.H., Boardman,L.A., Bugart,L.J., Burgess,C.C., Carino,T.J.,
Dwivedi,P., Huntress,M., Johnson,K.A., Lewis,M.E., Maitonis,P.J.,
Myerow,S.R., Brown-Shimer,S.L., Thilagalingam,A., Thibodeau,S.N. and
Molino,G.A.
Detection methods using TIMP 1 for colon cancer diagnosis
Patent: EP 1439393-A 53 21-JUL-2004;
JOURNAL Bayer Healthcare LLC (US); MAYO FOUNDATION FOR MEDICAL EDUCATION
AND RESEARCH (US)
FEATURES
source 1. 4139
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 60; DB 6; Length 4139;
Best Local Similarity 100.0%; Pred. No. 0.00076;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCTCCACCGGGCCCCCGAGCCACGGGTGTCACTTGGGCCCCGAGACACAGGCGCGCCCG 60

Db 458 GGCTCCACCGGGCCCCCGAGCCACGGGTGTCACTTGGGCCCCGAGACACAGGCGCGCCCG 517

Search completed: January 15, 2005, 22:36:42
Job time : 724.8 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2005, 22:37:00 : Search time 183.4 Seconds
(without alignments)
1879.789 Million cell updates/sec

Title: US-10-057-136-2

Perfect score: 60
Sequence: 1 ggcctaccgccccccagc.....cggacaccgccccgccc 60

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4300275 seqs, 2872944193 residues

Total number of hits satisfying chosen parameters: 8600550

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
21: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	60	100.0	60	US-10-057-136-2	Sequence 2, Appli
2	60	100.0	1424	US-10-447-839A-75	Sequence 75, Appl
3	60	100.0	1428	US-10-447-839A-20	Sequence 20, Appl
4	60	100.0	1527	US-10-057-136-19	Sequence 19, Appl
5	60	100.0	1799	US-10-447-839A-19	Sequence 19, Appl
6	60	100.0	1804	US-09-964-824A-573	Sequence 573, App
7	60	100.0	1804	US-10-029-517-17	Sequence 17, Appl
8	60	100.0	1804	US-10-717-920-84	Sequence 84, Appl
9	60	100.0	1823	US-10-101-510-339	Sequence 339, App
10	60	100.0	2297	US-10-406-317-41	Sequence 41, Appl
11	60	100.0	4139	US-09-964-824A-105	Sequence 105, App
12	60	100.0	4139	US-09-964-824A-105	Sequence 105, App

13	60	100.0	4139	US-09-964-824A-578	Sequence 578, App
14	60	100.0	4139	US-09-864-864-334	Sequence 334, App
15	60	100.0	4139	US-09-880-107-2121	Sequence 2121, App
16	60	100.0	4139	US-09-968-007A-751	Sequence 751, App
17	60	100.0	4139	US-10-171-311-157	Sequence 157, App
18	60	100.0	4139	US-10-177-293-310	Sequence 310, App
19	60	100.0	4139	US-10-440-464-155	Sequence 155, App
20	60	100.0	4139	US-10-734-564-53	Sequence 53, Appl
21	60	100.0	4139	US-10-775-920-80	Sequence 80, Appl
22	60	100.0	4139	US-10-775-920-85	Sequence 85, Appl
23	60	100.0	8181	US-10-447-839A-18	Sequence 18, Appl
24	60	100.0	8186	US-10-247-703-13	Sequence 23, Appl
25	60	100.0	8186	US-10-029-517-19	Sequence 19, Appl
26	58.4	97.3	572	US-10-029-517-18	Sequence 18, Appl
27	58.4	97.3	572	US-10-775-920-93	Sequence 93, Appl
28	58.4	97.3	1721	US-09-864-864-280	Sequence 280, App
29	58.4	97.3	1721	US-09-967-768A-224	Sequence 224, App
30	58.4	97.3	1721	US-10-247-703-21	Sequence 21, Appl
31	58.4	97.3	1721	US-10-097-340-211	Sequence 211, App
32	58.4	97.3	1721	US-10-171-311-155	Sequence 155, App
33	58.4	97.3	1721	US-10-007-926A-58	Sequence 58, Appl
34	58.4	97.3	1721	US-10-029-517-3	Sequence 3, Appl
35	58.4	97.3	1721	US-10-172-118-775	Sequence 775, App
36	58.4	97.3	1721	US-10-342-887-775	Sequence 775, App
37	58.4	97.3	1721	US-10-775-920-88	Sequence 88, Appl
38	58.4	97.3	2026	US-10-198-846-12589	Sequence 12589, A
39	58.4	97.3	2238	US-10-775-920-87	Sequence 87, Appl
40	58.4	97.3	2678	US-10-252-157-103	Sequence 103, App
41	57	95.0	518	US-10-247-703-38	Sequence 38, Appl
42	57	95.0	518	US-10-029-517-101	Sequence 101, App
43	56.8	94.7	78	US-10-057-136-13	Sequence 13, Appl
44	55.2	92.0	3343	US-10-247-703-27	Sequence 27, Appl
45	55.2	92.0	3343	US-10-029-517-102	Sequence 102, App

ALIGNMENTS

RESULT 1
US-10-057-136-2
Sequence 2, Application US/10057136
Publication No. US20030021770A1
GENERAL INFORMATION:
APPLICANT: SCHLOM, JEFFREY
APPLICANT: KANTOR, JUDITH
APPLICANT: KUFE, DONALD
APPLICANT: PANICALI, DENNIS
APPLICANT: GRITZ, LINDA
TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
FILE REFERENCE: 700953/47113C
CURRENT FILING DATE: 2002-01-25
PRIOR APPLICATION NUMBER: US/10/057,136
PRIOR FILING DATE: 1998-02-24
PRIOR APPLICATION NUMBER: 09/366,670
PRIOR FILING DATE: 1999-08-03
PRIOR APPLICATION NUMBER: PCT/US98/03693
PRIOR FILING DATE: 1998-02-24
PRIOR APPLICATION NUMBER: 60/038,253
PRIOR FILING DATE: 1997-02-24
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 60
TYPE: DNA
ORGANISM: Homo sapiens
US-10-057-136-2
Query Match 100.0% Score 60; DB 14; Length 60;
Best Local Similarity 100.0% Pred. No. 3.1e-09;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 ggcctaccgccccccagcgggtgacccgccccgagacacacgagccgcccgcg 60
|||||

Db 1 GGCTCCACCGCGCCCGCCAGCCCAAGGTGTACCTTCGGCCCGGACACCAAGCGCGCCCG 60

RESULT 2

US-10-447-839A-75/C
; Sequence 75, Application US/10447839A
; Publication No. US20040018181A1
; GENERAL INFORMATION:
; APPLICANT: Kufe, Donald W.
; APPLICANT: Kharbanda, Surender
; APPLICANT: Weltman, Steven D.
; TITLE OF INVENTION: MUC1 INTERFERENCE RNA COMPOSITIONS AND METHODS DERIVED THEREFROM
; FILE REFERENCE: 1000.05.009
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: US/10/447, 839A
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 09/951,938
; PRIOR FILING DATE: 2001-09-11
; PRIOR APPLICATION NUMBER: 60/231,841
; PRIOR FILING DATE: 2000-09-11
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 75
; LENGTH: 1424
; TYPE: RNA
; ORGANISM: ARTIFICIAL
; FEATURE:
; OTHER INFORMATION: Synthesized Sequence
US-10-447-839A-75

Query Match 100.0%; Score 60; DB 16; Length 1424;
Best Local Similarity 100.0%; Pred. No. 1.3e-09;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCTCCACCGCGCCCGCCAGCCCAAGGTGTACCTTCGGCCCGGACACCAAGCGCGCCCG 60
DB 1041 GGCTCCACCGCGCCCGCCAGCCCAAGGTGTACCTTCGGCCCGGACACCAAGCGCGCCCG 982

RESULT 3

US-10-447-839A-20
; Sequence 20, Application US/10447839A
; Publication No. US20040018181A1
; GENERAL INFORMATION:
; APPLICANT: Kufe, Donald W.
; APPLICANT: Kharbanda, Surender
; APPLICANT: Weltman, Steven D.
; TITLE OF INVENTION: MUC1 INTERFERENCE RNA COMPOSITIONS AND METHODS DERIVED THEREFROM
; FILE REFERENCE: 1000.05.009
; CURRENT APPLICATION NUMBER: US/10/447, 839A
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: 10/293,391
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 09/951,938
; PRIOR FILING DATE: 2001-09-11
; PRIOR APPLICATION NUMBER: 60/231,841
; PRIOR FILING DATE: 2000-09-11
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 20
; LENGTH: 1428
; TYPE: RNA
; ORGANISM: RNA
US-10-447-839A-20

Query Match 100.0%; Score 60; DB 16; Length 1428;
Best Local Similarity 93.3%; Pred. No. 1.3e-09;
Matches 56; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCTCCACCGCGCCCGCCAGCCCAAGGTGTACCTTCGGCCCGGACACCAAGCGCGCCCG 60
DB 385 GGCTCCACCGCGCCCGCCAGCCCAAGGTGTACCTTCGGCCCGGACACCAAGCGCGCCCG 444

RESULT 4

US-10-057-136-19
; Sequence 19, Application US/10057136
; Publication No. US20030021770A1
; GENERAL INFORMATION:
; APPLICANT: SCHLOM, JEFFREY
; APPLICANT: KANTOR, JUDITH
; APPLICANT: KUFEL, DONALD
; APPLICANT: PANICALI, DENNIS
; APPLICANT: GRITZ, LINDA
; TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1
; FILE REFERENCE: 700953/47113C
; CURRENT APPLICATION NUMBER: US/10/057,136
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/366,670
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: PCT/US98/03693
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/038,253
; PRIOR FILING DATE: 1997-02-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 19
; LENGTH: 1527
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-136-19

Query Match 100.0%; Score 60; DB 14; Length 1527;
Best Local Similarity 100.0%; Pred. No. 1.3e-09;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCTCCACCGCGCCCGCCAGCCCAAGGTGTACCTTCGGCCCGGACACCAAGCGCGCCCG 60
DB 226 GGCTCCACCGCGCCCGCCAGCCCAAGGTGTACCTTCGGCCCGGACACCAAGCGCGCCCG 285

RESULT 5

US-10-447-839A-19
; Sequence 19, Application US/10447839A
; Publication No. US20040018181A1
; GENERAL INFORMATION:
; APPLICANT: Kufe, Donald W.
; APPLICANT: Kharbanda, Surender
; APPLICANT: Weltman, Steven D.
; TITLE OF INVENTION: MUC1 INTERFERENCE RNA COMPOSITIONS AND METHODS DERIVED THEREFROM
; FILE REFERENCE: 1000.05.009
; CURRENT APPLICATION NUMBER: US/10/447, 839A
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: 10/293,391
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 09/951,938
; PRIOR FILING DATE: 2001-09-11
; PRIOR APPLICATION NUMBER: 60/231,841
; PRIOR FILING DATE: 2000-09-11
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 19
; LENGTH: 1799
; TYPE: RNA
; ORGANISM: RNA
US-10-447-839A-19

Query Match 100.0%; Score 60; DB 16; Length 1799;
Best Local Similarity 93.3%; Pred. No. 1.2e-09;
Matches 56; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCTCCACCGCGCCCGCCAGCCCAAGGTGTACCTTCGGCCCGGACACCAAGCGCGCCCG 60
DB 453 GGCTCCACCGCGCCCGCCAGCCCAAGGTGTACCTTCGGCCCGGACACCAAGCGCGCCCG 512

```
RESULT 6
US-09-964-824A-573
; Sequence 573, Application US/09964824A
; Patent No. US20020102531A1
; GENERAL INFORMATION:
; APPLICANT: Horrigan, Stephen
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-73
; CURRENT APPLICATION NUMBER: US/09/964, 824A
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/60/236, 033
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236, 032
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236, 028
; PRIOR FILING DATE: 2000-09-28
; NUMBER OF SEQ ID NOS: 583
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 573
; LENGTH: 1804
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-964-824A-573

Query Match          100.0%; Score 60; DB 9; Length 1804;
Best Local Similarity 100.0%; Pred. No. 1.2e-09;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCTCACCGCGCCCGCCAGCCGCGTGTACCTCGCGCCCGGACACCGAGCGCGCCCG 60
DB 457 GGCTCACCGCGCCCGCCAGCCGCGTGTACCTCGCGCCCGGACACCGAGCGCGCCCG 516

RESULT 7
US-10-029-517-17
; Sequence 17, Application US/10029517
; Publication No. US20030148969A1
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Susan J. Myers
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029, 517
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 17
; LENGTH: 1804
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (73)...(1500)
US-10-029-517-17

Query Match          100.0%; Score 60; DB 15; Length 1804;
Best Local Similarity 100.0%; Pred. No. 1.2e-09;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCTCACCGCGCCCGCCAGCCGCGTGTACCTCGCGCCCGGACACCGAGCGCGCCCG 60
DB 457 GGCTCACCGCGCCCGCCAGCCGCGTGTACCTCGCGCCCGGACACCGAGCGCGCCCG 516

RESULT 8
US-10-717-597-30
; Sequence 30, Application US/10717597
; Publication No. US20040110221A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Burczynski, Michael E.
```

```
; APPLICANT: Twine, Natalie C.
; APPLICANT: Dörner, Andrew J.
; APPLICANT: Trepicchio, William L.
; APPLICANT: Slonim, Donna K.
; APPLICANT: Stover, Jennifer A.
; TITLE OF INVENTION: METHODS FOR DIAGNOSING RCC AND OTHER SOLID TUMORS
; FILE REFERENCE: AM101080L
; CURRENT APPLICATION NUMBER: US/10/717, 597
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: US 60/459, 782
; PRIOR FILING DATE: 2003-04-03
; PRIOR APPLICATION NUMBER: US 60/427, 982
; PRIOR FILING DATE: 2002-11-21
; NUMBER OF SEQ ID NOS: 4904
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 30
; LENGTH: 1804
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-717-597-30

Query Match          100.0%; Score 60; DB 17; Length 1804;
Best Local Similarity 100.0%; Pred. No. 1.2e-09;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCTCACCGCGCCCGCCAGCCGCGTGTACCTCGCGCCCGGACACCGAGCGCGCCCG 60
DB 457 GGCTCACCGCGCCCGCCAGCCGCGTGTACCTCGCGCCCGGACACCGAGCGCGCCCG 516

RESULT 9
US-10-775-920-84
; Sequence 84, Application US/10775920
; Publication No. US2004017574A1
; GENERAL INFORMATION:
; APPLICANT: Merzen Ltd
; TITLE OF INVENTION: POLYNUCLEOTIDE SEQUENCES AND CORRESPONDING ENCODED POLYPEPTIDES
; TITLE OF INVENTION: OF PARTICULAR SECRETED AND MEMBRANE-BOUND PROTEINS OVEREXPRESSED
; FILE REFERENCE: Merzen - 0010B
; CURRENT APPLICATION NUMBER: US/10/775, 920
; CURRENT FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: US 60/447, 900
; PRIOR FILING DATE: 2003-02-13
; NUMBER OF SEQ ID NOS: 385
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 84
; LENGTH: 1804
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-775-920-84

Query Match          100.0%; Score 60; DB 17; Length 1804;
Best Local Similarity 100.0%; Pred. No. 1.2e-09;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCTCACCGCGCCCGCCAGCCGCGTGTACCTCGCGCCCGGACACCGAGCGCGCCCG 60
DB 457 GGCTCACCGCGCCCGCCAGCCGCGTGTACCTCGCGCCCGGACACCGAGCGCGCCCG 516

RESULT 10
US-10-101-510-339
; Sequence 339, Application US/10101510
; Publication No. US20030148295A1
; GENERAL INFORMATION:
; APPLICANT: WAN, JACKSON
; APPLICANT: WANG, YIXIN
; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
; FILE REFERENCE: 15117.0012
; CURRENT APPLICATION NUMBER: US/10/101, 510
; CURRENT FILING DATE: 2003-03-20
; PRIOR APPLICATION NUMBER: 60/276, 947
```

```
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 805
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 339
; LENGTH: 1823
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-101-510-339

Query Match
Best Local Similarity 100.0%; Score 60; DB 15; Length 1823;
Pred. No. 1.2e-09;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCTCCACCGCCCCCCCCAGCCCAAGGTGTCACTCGGCCCCGGACACCAAGGCGGCCCG 60
DB 457 GGCTCCACCGCCCCCCCCAGCCCAAGGTGTCACTCGGCCCCGGACACCAAGGCGGCCCG 516

RESULT 11
US-10-406-317-41
; Sequence 41, Application US/10406317
; Publication No. US20040019195A1
; GENERAL INFORMATION:
; APPLICANT: Schlom, Jeffrey;
; APPLICANT: Hodge, James;
; APPLICANT: Panicali, Dennis
; TITLE OF INVENTION: A recombinant vector expressing multiple constitutulatory
; FILE REFERENCE: 38163-0189
; CURRENT APPLICATION NUMBER: US/10/406,317
; CURRENT FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: US/09/856,988
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: PCT/US99/26866
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/111,582
; PRIOR FILING DATE: 1998-12-09
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 41
; LENGTH: 2297
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: VECTOR
US-10-406-317-41

Query Match
Best Local Similarity 100.0%; Score 60; DB 16; Length 2297;
Pred. No. 1.1e-09;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCTCCACCGCCCCCCCCAGCCCAAGGTGTCACTCGGCCCCGGACACCAAGGCGGCCCG 60
DB 346 GGCTCCACCGCCCCCCCCAGCCCAAGGTGTCACTCGGCCCCGGACACCAAGGCGGCCCG 405

RESULT 12
US-09-964-824A-105
; Sequence 105, Application US/09964824A
; Patent No. US20020102531A1
; GENERAL INFORMATION:
; APPLICANT: Horriqan, Stephen
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; FILE REFERENCE: 689290-73
; CURRENT APPLICATION NUMBER: US/09/964,824A
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/60/236,033
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,032
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,028
```

```
; PRIOR FILING DATE: 2000-09-28
; NUMBER OF SEQ ID NOS: 583
; SOFTWARE: Patent In Version 3.0
; SEQ ID NO 105
; LENGTH: 4139
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-964-824A-105

Query Match
Best Local Similarity 100.0%; Score 60; DB 9; Length 4139;
Pred. No. 9.7e-10;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCTCCACCGCCCCCCCCAGCCCAAGGTGTCACTCGGCCCCGGACACCAAGGCGGCCCG 60
DB 458 GGCTCCACCGCCCCCCCCAGCCCAAGGTGTCACTCGGCCCCGGACACCAAGGCGGCCCG 517

RESULT 13
US-09-964-824A-578
; Sequence 578, Application US/09964824A
; Patent No. US20020102531A1
; GENERAL INFORMATION:
; APPLICANT: Horriqan, Stephen
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; FILE REFERENCE: 689290-73
; CURRENT APPLICATION NUMBER: US/09/964,824A
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/60/236,033
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,032
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,028
; PRIOR FILING DATE: 2000-09-28
; NUMBER OF SEQ ID NOS: 583
; SOFTWARE: Patent In Version 3.0
; SEQ ID NO 578
; LENGTH: 4139
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-964-824A-578

Query Match
Best Local Similarity 100.0%; Score 60; DB 9; Length 4139;
Pred. No. 9.7e-10;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCTCCACCGCCCCCCCCAGCCCAAGGTGTCACTCGGCCCCGGACACCAAGGCGGCCCG 60
DB 458 GGCTCCACCGCCCCCCCCAGCCCAAGGTGTCACTCGGCCCCGGACACCAAGGCGGCCCG 517

RESULT 14
US-09-864-864-334
; Sequence 334, Application US/09864864
; Patent No. US20020102679A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Dillon, David C.
; APPLICANT: Secrist, Heather
; APPLICANT: Lodes, Michael J.
; APPLICANT: Algate, Paul A.
; APPLICANT: Fling, Steve P.
; APPLICANT: Mannion, Jane
; APPLICANT: Benson, Darrin R.
; APPLICANT: Carter, Darrick
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.523
; CURRENT APPLICATION NUMBER: US/09/864,864
; CURRENT FILING DATE: 2001-05-23
```


/ NUMBER OF SEQ ID NOS: 341
/ SOFTWARE: Corixa Invention Disclosure Database
/ SEQ ID NO 334
/ LENGTH: 4139
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-864-864-334

Query Match 100.0%; Score 60; DB 9; Length 4139;
Best Local Similarity 100.0%; Pred. No. 9.7e-10;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCTCCACCGCGCCCGCCAGCCAGCGGTGTCACTCGCGCCCGGACACAGCGCGGCCCG 60
|||
DB 458 GGCTCCACCGCGCCCGCCAGCCAGCGGTGTCACTCGCGCCCGGACACAGCGCGGCCCG 517

RESULT 15

US-09-880-107-2121
/ Sequence 2121, Application US/09880107
/ Patent No. US20020142981A1
/ GENERAL INFORMATION:
/ APPLICANT: Horne, Darci T.
/ APPLICANT: Vockley, Joseph G.
/ APPLICANT: Scherf, Uwe
/ APPLICANT: Gene Logic, Inc.
/ TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
/ FILE REFERENCE: 44921-5028-WO
/ CURRENT APPLICATION NUMBER: US/09/880,107
/ CURRENT FILING DATE: 2001-06-14
/ PRIOR APPLICATION NUMBER: US 60/211,379
/ PRIOR FILING DATE: 2000-06-14
/ PRIOR APPLICATION NUMBER: US 60/237,054
/ PRIOR FILING DATE: 2000-10-02
/ NUMBER OF SEQ ID NOS: 3950
/ SOFTWARE: Patent In Ver. 2.1
/ SEQ ID NO 2121
/ LENGTH: 4139
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: Genbank Accession No. US20020142981A1 J05582
US-09-880-107-2121

Query Match 100.0%; Score 60; DB 9; Length 4139;
Best Local Similarity 100.0%; Pred. No. 9.7e-10;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCTCCACCGCGCCCGCCAGCCAGCGGTGTCACTCGCGCCCGGACACAGCGCGGCCCG 60
|||
DB 458 GGCTCCACCGCGCCCGCCAGCCAGCGGTGTCACTCGCGCCCGGACACAGCGCGGCCCG 517

Search completed: January 16, 2005, 09:30:18
Job time : 184.4 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2005, 19:50:13 ; Search time 35.2 Seconds
(without alignments)
1211.572 Million cell updates/sec

Title: US-10-057-136-2

Perfect score: 60

Sequence: 1 ggcctcaccgccccccagc.....cggacaccgagccgccccg 60

Scoring table:

IDENTITY NUC

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/prodata/1/ina/5A_COMB.seq: *
2: /cgn2_6/prodata/1/ina/5B_COMB.seq: *
3: /cgn2_6/prodata/1/ina/6A_COMB.seq: *
4: /cgn2_6/prodata/1/ina/6B_COMB.seq: *
5: /cgn2_6/prodata/1/ina/PCTUS_COMB.seq: *
6: /cgn2_6/prodata/1/ina/Backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	1804	4	US-10-029-517-17 Sequence 17, Appl
2	60	100.0	8186	4	US-10-029-517-19 Sequence 19, Appl
3	58.4	97.3	572	4	US-10-029-517-18 Sequence 18, Appl
4	58.4	97.3	1721	4	US-10-029-517-3 Sequence 3, Appl
5	57	95.0	518	4	US-10-029-517-101 Sequence 101, App
6	55.2	92.0	3343	4	US-10-029-517-102 Sequence 102, App
7	54	90.0	60	4	US-09-475-947A-246 Sequence 246, App
8	52.2	87.0	981	4	US-10-029-517-16 Sequence 16, Appl
9	51	85.0	6192	2	US-08-479-537A-1 Sequence 1, Appl
10	51	85.0	6192	3	US-09-083-116-1 Sequence 1, Appl
11	51	85.0	6192	3	US-09-134-916A-1 Sequence 1, Appl
12	51	85.0	6449	2	US-08-479-537A-4 Sequence 4, Appl
13	51	85.0	6449	3	US-09-083-116-4 Sequence 4, Appl
14	51	85.0	6449	3	US-09-134-916A-4 Sequence 4, Appl
15	29.6	49.3	556	4	US-10-029-517-105 Sequence 105, App
16	27.6	46.0	5577	1	US-08-326-117B-1 Sequence 1, Appl
17	27.6	46.0	5577	3	US-08-982-129-1 Sequence 1, Appl
18	27.6	46.0	5582	4	US-09-178-176B-1 Sequence 1, Appl
19	27.6	46.0	5582	4	US-09-457-864-1 Sequence 1, Appl
20	27.6	46.0	5582	4	US-09-457-865A-1 Sequence 1, Appl
21	27.2	45.3	1441	1	US-08-136-277-18 Sequence 18, Appl
22	27.2	45.3	1441	2	US-08-479-403-18 Sequence 18, Appl
23	27.2	45.3	1441	2	US-08-835-734-18 Sequence 18, Appl
24	27.2	45.3	2750	1	US-08-136-277-1 Sequence 1, Appl
25	27.2	45.3	2750	2	US-08-479-403-1 Sequence 1, Appl
26	27.2	45.3	2750	3	US-08-835-734-1 Sequence 1, Appl
27	27.2	45.3	77536	4	US-09-410-551B-1 Sequence 1, Appl

28	27.2	45.3	77536	4	US-09-940-316B-1 Sequence 1, Appl
29	27	45.0	429	4	US-09-854-133-713 Sequence 713, App
30	26.8	44.7	513	4	US-09-646-028-42 Sequence 42, Appl
31	26.8	44.7	534	4	US-09-646-028-46 Sequence 46, Appl
32	26.6	44.3	1737	2	US-08-750-703-2 Sequence 2, Appl
33	26.6	44.3	8779	2	US-08-750-703-4 Sequence 4, Appl
34	26.4	44.0	774	3	US-08-415-655-11 Sequence 11, Appl
35	26.4	44.0	3245	4	US-09-774-528-352 Sequence 352, App
36	26.4	44.0	4480	3	US-09-191-171-7 Sequence 7, Appl
37	26.4	44.0	4480	3	US-09-385-707-7 Sequence 7, Appl
38	26.4	44.0	6238	4	US-09-639-696C-6 Sequence 6, Appl
39	26.4	44.0	8438	1	US-07-945-283-1 Sequence 1, Appl
40	26.4	44.0	15378	3	US-08-785-420-1 Sequence 1, Appl
41	26	43.3	2581	3	US-09-363-708-1 Sequence 1, Appl
42	26	43.3	2581	4	US-09-083-587-1 Sequence 1, Appl
43	26	43.3	2743	1	US-08-396-479B-3 Sequence 3, Appl
44	26	43.3	2743	1	US-08-818-823-3 Sequence 3, Appl
45	26	43.3	2749	2	US-08-124-981A-1 Sequence 1, Appl

ALIGNMENTS

```
RESULT 1
US-10-029-517-17
; Sequence 17, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 17
; LENGTH: 1804
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (73)...(1500)
US-10-029-517-17

Query Match          100.0%; Score 60; DB 4; Length 1804;
Best Local Similarity 100.0%; Pred. No. 6e-08;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCCTCACCGCCGCCGACGAGTGTGACCTGCGCCGGACACCGAGCGGCCCG 60
DB 457 GGCCTCACCGCCGCCGACGAGTGTGACCTGCGCCGGACACCGAGCGGCCCG 516

RESULT 2
US-10-029-517-19
; Sequence 19, Application US/10029517
; Patent No. 6716627
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
; FILE REFERENCE: RTS-0352
; CURRENT APPLICATION NUMBER: US/10/029,517
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 107
; SEQ ID NO 19
; LENGTH: 8186
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 6899
; OTHER INFORMATION: unknown
```

```
/ NAME/KEY: unsure
/ LOCATION: 7155
/ OTHER INFORMATION: unknown
/ NAME/KEY: unsure
/ LOCATION: 7184
/ OTHER INFORMATION: unknown
/ NAME/KEY: unsure
/ LOCATION: 7957
/ OTHER INFORMATION: unknown
/ NAME/KEY: intron
/ LOCATION: (2997)...(3498)
/ OTHER INFORMATION: intron 1
/ NAME/KEY: intron:exon junction
/ LOCATION: (3498)...(3499)
/ OTHER INFORMATION: intron 1:exon 2
/ NAME/KEY: exon
/ LOCATION: (3508)...(3599)
/ OTHER INFORMATION: exon 2d
/ NAME/KEY: exon:intron junction
/ LOCATION: (3982)...(3983)
/ OTHER INFORMATION: exon 2a:intron 2a
/ NAME/KEY: intron:exon junction
/ LOCATION: (4205)...(4206)
/ OTHER INFORMATION: intron 2c:exon 3c
/ NAME/KEY: intron:exon junction
/ LOCATION: (4259)...(4260)
/ OTHER INFORMATION: intron 2d:exon 3d
/ NAME/KEY: exon
/ LOCATION: (4260)...(4328)
/ OTHER INFORMATION: exon 3d
/ NAME/KEY: intron:exon junction
/ LOCATION: (4632)...(4633)
/ OTHER INFORMATION: intron 3:exon 4
/ NAME/KEY: exon
/ LOCATION: (4914)...(5035)
/ OTHER INFORMATION: exon 5
/ NAME/KEY: intron
/ LOCATION: (5286)...(6293)
/ OTHER INFORMATION: intron 6
/ US-10-029-517-19
```

```
Query Match
Best Local Similarity 100.0%; Score 60; DB 4; Length 8186;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 GGCTCCACCGCCCCCGGAGCCGAGGTGTCACTCGGCGCCGAGACACGAGCGGCGCCG 60
Db 3825 GGCTCCACCGCCCCCGGAGCCGAGGTGTCACTCGGCGCCGAGACACGAGCGGCGCCG 3884
```

```
RESULT 3
US-10-029-517-18
/ Sequence 18, Application US/10029517
/ Patent No. 6716627
/ GENERAL INFORMATION:
/ APPLICANT: Kenneth W. Dobie
/ TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
/ FILE REFERENCE: RTS-0352
/ CURRENT APPLICATION NUMBER: US/10/029,517
/ NUMBER OF SEQ ID NOS: 107
/ SEQ ID NO 18
/ LENGTH: 572
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (67)...(572)
/ US-10-029-517-18
```

```
Query Match
Best Local Similarity 97.3%; Score 58.4; DB 4; Length 572;
Matches 57; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GGCTCCACCGCCCCCGGAGCCGAGGTGTCACTCGGCGCCGAGACACGAGCGGCGCCG 60
Db 478 GGCTCCACCGCCCCCGGAGCCGAGGTGTCACTCGGCGCCGAGACACGAGCGGCGCCG 537
```

```
RESULT 4
US-10-029-517-3
/ Sequence 3, Application US/10029517
/ Patent No. 6716627
/ GENERAL INFORMATION:
/ APPLICANT: Kenneth W. Dobie
/ TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
/ FILE REFERENCE: RTS-0352
/ CURRENT APPLICATION NUMBER: US/10/029,517
/ NUMBER OF SEQ ID NOS: 107
/ SEQ ID NO 3
/ LENGTH: 1721
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (58)...(1605)
/ US-10-029-517-3
```

```
Query Match
Best Local Similarity 97.3%; Score 58.4; DB 4; Length 1721;
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 GGCTCCACCGCCCCCGGAGCCGAGGTGTCACTCGGCGCCGAGACACGAGCGGCGCCG 60
Db 442 GGCTCCACCGCCCCCGGAGCCGAGGTGTCACTCGGCGCCGAGACACGAGCGGCGCCG 501
```

```
RESULT 5
US-10-029-517-101
/ Sequence 101, Application US/10029517
/ Patent No. 6716627
/ GENERAL INFORMATION:
/ APPLICANT: Kenneth W. Dobie
/ TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
/ FILE REFERENCE: RTS-0352
/ CURRENT APPLICATION NUMBER: US/10/029,517
/ NUMBER OF SEQ ID NOS: 107
/ SEQ ID NO 101
/ LENGTH: 518
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ US-10-029-517-101
```

```
Query Match
Best Local Similarity 95.0%; Score 57; DB 4; Length 518;
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 GGCTCCACCGCCCCCGGAGCCGAGGTGTCACTCGGCGCCGAGACACGAGCGGCGCC 57
Db 462 GGCTCCACCGCCCCCGGAGCCGAGGTGTCACTCGGCGCCGAGACACGAGCGGCGCC 518
```

```
RESULT 6
US-10-029-517-102
/ Sequence 102, Application US/10029517
/ Patent No. 6716627
/ GENERAL INFORMATION:
/ APPLICANT: Kenneth W. Dobie
/ TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
```

FILE REFERENCE: RTS-0352
CURRENT APPLICATION NUMBER: US/10/029,517
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 107
SEQ ID NO 102
LENGTH: 3343
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
US-10-029-517-102

Query Match 92.0%; Score 55.2; DB 4; Length 3343;
Best Local Similarity 95.0%; Pred. No. 1.1e-06;
Matches 57; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GGCTCCACCGCCCCCGCCAGCCCGGTGTCACTCGGCCCGGACACCGCCGCGCCG 60
Db 1668 GGCTCCACCGCCCCCGCCAGCCCGGTGTCACTCGGCCCGGAGACACCGCCGCGCCG 1727

RESULT 7
US-09-475-947A-246
Sequence 246, Application US/09475947A
Patent No. 6472154
GENERAL INFORMATION:
APPLICANT: Garner, Harold R.
APPLICANT: Wren, Jonathan D.
FILE REFERENCE: Polymorphic Repeats in Human Genes
FILE REFERENCE: UTS0667
CURRENT APPLICATION NUMBER: US/09/475,947A
CURRENT FILING DATE: 1999-12-31
NUMBER OF SEQ ID NOS: 346
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 246
LENGTH: 60
TYPE: DNA
ORGANISM: human
US-09-475-947A-246

Query Match 90.0%; Score 54; DB 4; Length 60;
Best Local Similarity 100.0%; Pred. No. 2.8e-06;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCTCCACCGCCCCCGCCAGCCCGGTGTCACTCGGCCCGGACACCGCCGCG 54
Db 7 GGCTCCACCGCCCCCGCCAGCCCGGTGTCACTCGGCCCGGACACCGCCGCG 60

RESULT 8
US-10-029-517-16
Sequence 16, Application US/10029517
Patent No. 6716627
GENERAL INFORMATION:
APPLICANT: Kenneth W. Dobie
APPLICANT: Susan J. Myers
TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
FILE REFERENCE: RTS-0352
CURRENT APPLICATION NUMBER: US/10/029,517
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 107
SEQ ID NO 16
LENGTH: 981
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: exon:exon junction
LOCATION: (464)...(465)
OTHER INFORMATION: exon 3b:exon 4
US-10-029-517-16

Query Match 87.0%; Score 52.2; DB 4; Length 981;
Best Local Similarity 94.7%; Pred. No. 7.3e-06;

Matches 54; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 GGCTCCACCGCCCCCGCCAGCCCGGTGTCACTCGGCCCGGACACCGCCGCGCCG 57
Db 21 GGCTCCACCGCCCCCGCCAGCCCGGTGTCACTCGGCCCGGAGACACCGCCGCGCCG 77

RESULT 9
US-08-479-537A-1
Sequence 1, Application US/08479537A
Patent No. 5861381
GENERAL INFORMATION:
APPLICANT: CHAMBER, Pierre
APPLICANT: KIENY, Marie-Paule
APPLICANT: LATHE, Richard
APPLICANT: HAREVENT, Mara
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
TREATMENT OR PREVENTION OF A MALIGNANT TUMOR
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/479,537A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 90/13101
FILING DATE: 23-OCT-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/FR91/00835
FILING DATE: 23-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/039,320
FILING DATE: 04-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/403,576
FILING DATE: 14-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Teskin, Robin L.
REGISTRATION NUMBER: 35,030
REFERENCE/DOCKET NUMBER: 017753-025
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 6192 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 58..120
FEATURE:
NAME/KEY: repeat_region
LOCATION: 439..5239
OTHER INFORMATION: /note="The nucleotides spanning 439-5239 constitute a repeated region wherein the repeat is a nucleotide and encodes 20 amino acids, 17 of which are fixed
OTHER INFORMATION: The number of such repeats varies from 1 to 80."
FEATURE:
NAME/KEY: mat_peptide

LOCATION: 121..6166
FEATURE:
NAME/KEY: repeat_region
LOCATION: 457
OTHER INFORMATION: /note= "Nucleotide 457 is X1 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 487
OTHER INFORMATION: /note= "Nucleotide 487 is Y = NNN
OTHER INFORMATION: which is the codon for Thr or Asn wherein Thr = ACT, ACC, ACA
OTHER INFORMATION: or ACG; and Asn = AAT or AAC."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 496
OTHER INFORMATION: /note= "Nucleotide 496 is X2 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
US-08-479-537A-1

Query Match 85.0%; Score 51; DB 2; Length 6192;
Best Local Similarity 85.0%; Pred. No. 1,4e-05;
Matches 51; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 GGCTCCACGGCCCCCGACGCGGTGTCACCTCGGCGCCCGGACACACGCGCGCCCG 60
Db 442 GGCTCCACGGCCCCCGACGCGGTGTCACCTCGGCGCCCGGACACGCGCGCCCG 501

RESULT 10
US-09-083-116-1
Sequence 1, Application US/09083116
Patent No. 6203795

GENERAL INFORMATION:
APPLICANT: CHAMON, Pierre
APPLICANT: KIENY, Marie-Paule
APPLICANT: LATHE, Richard
APPLICANT: HAREUVENI, Mara
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRES:
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/083,116
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/479,537
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/FR91/00835
FILING DATE: 23-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/039,320
FILING DATE: 04-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/403,576
FILING DATE: 14-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Teskin, Robin L.
REGISTRATION NUMBER: 35,030

REFERENCE/DOCKET NUMBER: 017753-025
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 6192 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 58..120
FEATURE:
NAME/KEY: repeat_region
LOCATION: 439..5239
OTHER INFORMATION: /note= "The nucleotides spanning
OTHER INFORMATION: 439-5239 constitute a repeated region wherein the repeat is 6
OTHER INFORMATION: nucleotides and encodes 20 amino acids, 17 of which are fixed
OTHER INFORMATION: The number of such repeats varies from 1 to 80."
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 121..6166
FEATURE:
NAME/KEY: repeat_region
LOCATION: 457
OTHER INFORMATION: /note= "Nucleotide 457 is X1 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 487
OTHER INFORMATION: /note= "Nucleotide 487 is Y = NNN
OTHER INFORMATION: which is the codon for Thr or Asn wherein Thr = ACT, ACC, ACA
OTHER INFORMATION: or ACG; and Asn = AAT or AAC."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 496
OTHER INFORMATION: /note= "Nucleotide 496 is X2 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
US-09-083-116-1

Query Match 85.0%; Score 51; DB 3; Length 6192;
Best Local Similarity 85.0%; Pred. No. 1,4e-05;
Matches 51; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 GGCTCCACGGCCCCCGACGCGGTGTCACCTCGGCGCCCGGACACGCGCGCCCG 60
Db 442 GGCTCCACGGCCCCCGACGCGGTGTCACCTCGGCGCCCGGACACGCGCGCCCG 501

RESULT 11
US-09-134-916A-1
Sequence 1, Application US/09134916A
Patent No. 6328956

GENERAL INFORMATION:
APPLICANT: CHAMON, Pierre
APPLICANT: KIENY, Marie-Paule
APPLICANT: LATHE, Richard
APPLICANT: HAREUVENI, Mara
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRES:
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/134,916A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/479,537
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: FR 90/13101
FILING DATE: 23-OCT-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/FR91/00835
FILING DATE: 23-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/039,320
FILING DATE: 04-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/403,576
FILING DATE: 14-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Teskin, Robin L.
REGISTRATION NUMBER: 35,030
REFERENCE/DOCKET NUMBER: 017753-025
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 6192 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 58..120
FEATURE:
NAME/KEY: repeat_region
LOCATION: 439..5239
OTHER INFORMATION: /note="The nucleotides spanning 439-5239 constitute a repeated region wherein the repeat is 6
OTHER INFORMATION: nucleotides and encodes 20 amino acids, 17 of which are fixed
OTHER INFORMATION: The number of such repeats varies from 1 to 80."
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 121..6166
FEATURE:
NAME/KEY: repeat_region
LOCATION: 457
OTHER INFORMATION: /note="Nucleotide 457 is X1 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 487
OTHER INFORMATION: /note="Nucleotide 487 is Y = NNN
OTHER INFORMATION: which is the codon for Thr or Asn wherein Thr = ACT, ACC, ACA
OTHER INFORMATION: or ACG; and Asn = AAT or AAC."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 496
OTHER INFORMATION: /note="Nucleotide 496 is X2 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
US-09-134-916A-1
Query Match 85.0%; Score 51; DB 3; Length 6192;
Best Local Similarity 85.0%; Pred. No. 1.4e-05;
Matches 51; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 GGCTCACCAGCCGCCCCCAGGAGTGTACCTCGAGCCCGAGACAGACGCGGCCG 60
|||||
Db 442 GGCTCACCAGCCGCCCCCAGGAGTGTACCTCGAGCCCGAGACAGGCGCCG 501
|||||
RESULT 12
US-08-479-537A-4
Sequence 4, Application US/08479537A
Patent No. 5861381
GENERAL INFORMATION:
APPLICANT: CHAMMON, Pierre
APPLICANT: KIENY, Marie-Paule
APPLICANT: LATHE, Richard
APPLICANT: HAREIVENT, Mara
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
TREATMENT OR PREVENTION OF A MALIGNANT TUMOR
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/479,537A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 90/13101
FILING DATE: 23-OCT-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/FR91/00835
FILING DATE: 23-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/039,320
FILING DATE: 04-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/403,576
FILING DATE: 14-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Teskin, Robin L.
REGISTRATION NUMBER: 35,030
REFERENCE/DOCKET NUMBER: 017753-025
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 6449 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 58..120
FEATURE:
NAME/KEY: repeat_region
LOCATION: 439..5239
OTHER INFORMATION: /note="The nucleotides spanning 439-5239 constitute a repeated region wherein the repeat is 6
OTHER INFORMATION: nucleotides and encodes 20 amino acids, 17 of which are fixed
OTHER INFORMATION: The number of such repeats varies from 1 to 80."
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 121..5661
FEATURE:

NAME/KEY: repeat_region
LOCATION: 457
OTHER INFORMATION: /note= "Nucleotide 457 is X1 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 487
OTHER INFORMATION: /note= "Nucleotide 487 is Y = NNN
OTHER INFORMATION: which is the codon for Thr or Asn wherein Thr = ACT, ACC, ACA
OTHER INFORMATION: or ACG; and Asn = AAT or AAC."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 496
OTHER INFORMATION: /note= "Nucleotide 496 is X2 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
US-08-479-537A-4

Query Match
Best Local Similarity 85.0%; Score 51; DB 2; Length 6449;
Matches 51; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 GGCTCCACCGCCCCCGCCAGCGGTGTCACTCGGCCCCCGAGACACGAGCGGCCCG 60
Db 442 GGCTCCACCGCCCCCGCCAGCGGTGTCACTCGGCCCCCGAGACACGAGCGGCCCG 501

RESULT 13
US-09-083-116-4

Sequence 4, Application US/09083116
Patent No. 6203795

GENERAL INFORMATION:
APPLICANT: CHAMON, Pierre
APPLICANT: KIENY, Marie-Paule
APPLICANT: LATHE, Richard
APPLICANT: HAREVIENT, Mara
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESS: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/083,116
FILING DATE:

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/479,537
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/FR91/00835
FILING DATE: 23-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/039,320
FILING DATE: 04-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/403,576
FILING DATE: 14-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Teskin, Robin L.
REGISTRATION NUMBER: 35,030
REFERENCE/DOCKET NUMBER: 017753-025
TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 6449 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 58..120
FEATURE:
NAME/KEY: repeat_region
LOCATION: 439..5239
OTHER INFORMATION: /note= "The nucleotides spanning
OTHER INFORMATION: 439-5239 constitute a repeated region wherein the repeat is 6
OTHER INFORMATION: nucleotides and encodes 20 amino acids, 17 of which are fixed
OTHER INFORMATION: The number of such repeats varies from 1 to 80."
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 121..5661
FEATURE:
NAME/KEY: repeat_region
LOCATION: 457
OTHER INFORMATION: /note= "Nucleotide 457 is X1 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 487
OTHER INFORMATION: /note= "Nucleotide 487 is Y = NNN
OTHER INFORMATION: which is the codon for Thr or Asn wherein Thr = ACT, ACC, ACA
OTHER INFORMATION: or ACG; and Asn = AAT or AAC."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 496
OTHER INFORMATION: /note= "Nucleotide 496 is X2 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
US-09-083-116-4

Query Match
Best Local Similarity 85.0%; Score 51; DB 3; Length 6449;
Matches 51; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 GGCTCCACCGCCCCCGCCAGCGGTGTCACTCGGCCCCCGAGACACGAGCGGCCCG 60
Db 442 GGCTCCACCGCCCCCGCCAGCGGTGTCACTCGGCCCCCGAGACACGAGCGGCCCG 501

RESULT 14
US-09-134-916A-4

Sequence 4, Application US/09134916A
Patent No. 6328956

GENERAL INFORMATION:
APPLICANT: CHAMON, Pierre
APPLICANT: KIENY, Marie-Paule
APPLICANT: LATHE, Richard
APPLICANT: HAREVIENT, Mara
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESS: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 6449 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 58..120
FEATURE:
NAME/KEY: repeat_region
LOCATION: 439..5239
OTHER INFORMATION: /note= "The nucleotides spanning
OTHER INFORMATION: 439-5239 constitute a repeated region wherein the repeat is 6
OTHER INFORMATION: nucleotides and encodes 20 amino acids, 17 of which are fixed
OTHER INFORMATION: The number of such repeats varies from 1 to 80."
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 121..5661
FEATURE:
NAME/KEY: repeat_region
LOCATION: 457
OTHER INFORMATION: /note= "Nucleotide 457 is X1 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 487
OTHER INFORMATION: /note= "Nucleotide 487 is Y = NNN
OTHER INFORMATION: which is the codon for Thr or Asn wherein Thr = ACT, ACC, ACA
OTHER INFORMATION: or ACG; and Asn = AAT or AAC."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 496
OTHER INFORMATION: /note= "Nucleotide 496 is X2 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
US-09-083-116-4

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/134, 916A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/479, 537
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: FR 90/13101
FILING DATE: 23-OCT-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/FR91/00835
FILING DATE: 23-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/039,320
FILING DATE: 04-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/403,576
FILING DATE: 14-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Teskin, Robin L.
REGISTRATION NUMBER: 35,030
REFERENCE/DOCKET NUMBER: 017753-025
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 6449 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 58..120
FEATURE:
NAME/KEY: repeat_region
LOCATION: 439..5239
OTHER INFORMATION: /note= "The nucleotides spanning 439-5239 constitute a repeated region wherein the repeat is 6
OTHER INFORMATION: nucleotides and encodes 20 amino acids, 17 of which are fixed
OTHER INFORMATION: The number of such repeats varies from 1 to 80."
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 121..5661
FEATURE:
NAME/KEY: repeat_region
LOCATION: 457
OTHER INFORMATION: /note= "Nucleotide 457 is X1 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 487
OTHER INFORMATION: /note= "Nucleotide 487 is Y = NNN
OTHER INFORMATION: which is the codon for Thr or Asn wherein Thr = ACT, ACC, ACA
OTHER INFORMATION: or ACG; and Asn = AAT or AAC."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 496
OTHER INFORMATION: /note= "Nucleotide 496 is X2 = NNN
OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro = CCT, CCC, CCA
OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG."
US-09-134-916A-4
Query Match 85.0%; Score 51; DB 3; Length 6449;
Best Local Similarity 85.0%; Pred. No. 1.4e-05;
Matches 51; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

1 GGCTCCACCGCCCGCCGAGGTGTACCTCGCGCCGAGACACGAGCCGCCCG 60
|||||

Db 442 GGCTCCACCGCCCGCCGAGGTGTACCTCGCGCCGAGACACGAGCCGCCCG 501
RESULT 15
US-10-029-517-105
Sequence 105, Application US/10029517
Patent No. 6716627
GENERAL INFORMATION:
APPLICANT: Kenneth W. Doble
APPLICANT: Susan J. Myers
TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION
FILE REFERENCE: RTS-0352
CURRENT APPLICATION NUMBER: US/10/029,517
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 107
SEQ ID NO 105
LENGTH: 556
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: 5
OTHER INFORMATION: n = A,T,C or G
US-10-029-517-105

Query Match 49.3%; Score 29.6; DB 4; Length 556;
Best Local Similarity 68.3%; Pred. No. 7.4;
Matches 41; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
Db 481 GGTTCAAGCTGCACCTGGGAGACGATGTCACTTGTCCAGTACACGAGCCGCTG 540
|||||

Search completed: January 16, 2005, 03:01:14
Job time : 36.2 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd

OM nucleic - nucleic search, using sw model

```
Run on:      January 15, 2005, 16:00:11 ; Search time 170.5 Seconds
              (without alignments)
              1847.304 Million cell updates/sec
```

Title:	US-10-057-136-2
Perfect score:	60
Sequence:	1 ggtccacgcgccccccagc.....cgacaccaggccggcccgcg 60

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4134886 segs, 2624710521 residues
Total number of hits satisfying chosen parameters: 8269772

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

```
Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

Database : N_GeneSeq_23Sep04: *

1: N_genseqqr_235ep04: *
2: genseeqqr1990s: *
3: genseeqqr2000as: *
4: genseeqqr2001as: *
5: genseeqqr2002as: *
6: genseeqqr2002as: *
7: genseeqqr2002as: *
8: genseeqqr2003as: *
9: genseeqqr2003as: *
10: genseeqqr2003as: *
11: genseeqqr2003as: *
12: genseeqqr2004as: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	60	100.0	60	2	AAV48316	AAV48316 Nucleotid
c 2	60	100.0	309	1	AAAN90579	AAAN90579 pDf9.3 cD
3	60	100.0	1194	12	AD1577112	AD1577112 Human brc
4	60	100.0	1378	12	AD157693	AD157693 Human brc
c 5	60	100.0	1424	12	AD023180	AD023180 Artisenso
6	60	100.0	1428	6	AB160159	AB160159 Human MUC
7	60	100.0	1428	12	AD023125	AD023125 Human MUC
8	60	100.0	1457	12	ADF32627	ADF32627 Plasmid c
9	60	100.0	1527	2	AAV48319	AAV48319 MiniMUC1
10	60	100.0	1614	12	ADK70370	ADK70370 Respirator
11	60	100.0	1630	12	AD157708	AD157708 Human brc
12	60	100.0	1634	12	AD157689	AD157689 Human brc
13	60	100.0	1712	12	AD157686	AD157686 Human brc
14	60	100.0	1728	12	AD157669	AD157669 Human brc
15	60	100.0	1755	12	AD157673	AD157673 Human brc
16	60	100.0	1774	12	AD043991	AD043991 Plasmid c
17	60	100.0	1774	12	ADF32625	ADF32625 Plasmid c
18	60	100.0	1799	12	AD023124	AD023124 Human MUC
19	60	100.0	1803	12	AD157699	AD157699 Human brc
20	60	100.0	1804	6	AB167539	AB167539 Thyroid c
21	60	100.0	1804	9	AA056950	AA056950 Human muc

45	60	100.0	4139	6	Ab167071	Thyroid c
44	60	100.0	2333	12	Ad157665	Human bre
43	60	100.0	2255	12	Ad157667	Human bre
42	60	100.0	2194	12	Ad157663	Human bre
41	60	100.0	2135	12	AdF32629	Plaemid J
40	60	100.0	2094	12	Ad157661	Human bre
39	60	100.0	2049	12	Ad157662	Human bre
38	60	100.0	2048	12	Ad157668	Human bre
37	60	100.0	2045	12	Ad157701	Human bre
36	60	100.0	1953	12	Ad157668	Human bre
35	60	100.0	1949	12	Ad157659	Human bre
34	60	100.0	1945	12	Ad157676	Human bre
33	60	100.0	1930	12	Ad157678	Human bre
32	60	100.0	1918	12	Ad157672	Human bre
31	60	100.0	1892	12	Ad157677	Human bre
30	60	100.0	1874	12	Ad157668	Human bre
29	60	100.0	1835	12	AdF32631	Plaemid J
28	60	100.0	1823	6	Ab235228	Human gen
27	60	100.0	1823	12	Ad157707	Human bre
26	60	100.0	1818	12	AdF32633	Plaemid J
25	60	100.0	1808	12	Ad157706	Human bre
24	60	100.0	1805	12	AdO28642	Human MIO
23	60	100.0	1804	10	Adp13294	Renal cel
22	60	100.0	1804	10	AdD14719	Human arc
21	60	100.0	1804	12	Adp13294	Renal cel
20	60	100.0	1804	12	Adp13294	Renal cel
19	60	100.0	1804	12	Adp13294	Renal cel
18	60	100.0	1804	12	Adp13294	Renal cel
17	60	100.0	1804	12	Adp13294	Renal cel
16	60	100.0	1804	12	Adp13294	Renal cel
15	60	100.0	1804	12	Adp13294	Renal cel
14	60	100.0	1804	12	Adp13294	Renal cel
13	60	100.0	1804	12	Adp13294	Renal cel
12	60	100.0	1804	12	Adp13294	Renal cel
11	60	100.0	1804	12	Adp13294	Renal cel
10	60	100.0	1804	12	Adp13294	Renal cel
9	60	100.0	1804	12	Adp13294	Renal cel
8	60	100.0	1804	12	Adp13294	Renal cel
7	60	100.0	1804	12	Adp13294	Renal cel
6	60	100.0	1804	12	Adp13294	Renal cel
5	60	100.0	1804	12	Adp13294	Renal cel
4	60	100.0	1804	12	Adp13294	Renal cel
3	60	100.0	1804	12	Adp13294	Renal cel
2	60	100.0	1804	12	Adp13294	Renal cel
1	60	100.0	1804	12	Adp13294	Renal cel

ALIGNMENTS

RESULT 1
AAV48316

DT 20-NOV-1998 (first entry)

Nucleotide sequence encoding MUC1 tandem repeat unit

KW 88; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen.
KW tumour; tumour-associated antigen.

OS Homo sapiens.

Key	Location/Qualifiers
FH	1. .60
FT	/*tag= a
FT	/product= "MUCl tandem repeat unit
FT	

PN W09837095-A2.

PD 27-AUG-1998.

PF 24-FEB-1998; 98WO-US003693

PR 24-FEB-1997; 97US-0038253P

PA (THER-) THERION BIOLOGICS CORP.

PA (DAND) DANA FARBER CANCER INST INC

PI Schlom J, Kantor J, Kufe D, Panicali D, Gritz L,

DR WPI; 1998-467492/40.

XX
DE

PT immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.

Example 1; page 20; 42pp; English.

CC The MUC1 tandem repeat unit was used to create an immunogenic mini-MUC1
CC fragment for inclusion in a recombinant pox virus (RPV). The RPV was used

CC in a pharmaceutical composition also containing an immunomodulator to
 CC generate MUC1 specific cytotoxic T-lymphocytes. The recombinant pox virus
 CC therefore encodes an immunogenic MUC1 fragment that does not undergo
 CC significant genetic deletion, thereby providing an unexpectedly stable
 CC and immunogenic pox virus. They can be used to prevent or treat tumours
 CC expressing MUC1 tumour-associated antigens

Sequence 60 BP; 7 A; 33 C; 16 G; 4 T; 0 U; 0 Other;

Query Match 100.0%; Score 60; DB 2; Length 60;
 Best Local Similarity 100.0%; Pred. No. 1.8e-06;
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCTCCACCGCCCGCCCGCCGCGGTGTACCTCGGCGCCGACACCGAGCGCGCCCG 60
 DB 1 GGCTCCACCGCCCGCCCGCCGCGGTGTACCTCGGCGCCGACACCGAGCGCGCCCG 60

RESULT 2

AA90579/c
 ID AA90579 standard; cDNA; 309 BP.

AC AA90579;
 DT 27-AUG-2003 (revised)
 DT 25-MAR-2003 (revised)
 DT 04-DEC-1989 (first entry)

DE PDF9.3 cDNA insert.

KW PDF9.3; human DF3 breast carcinoma-associated antigen epitope.

OS Human MCF-7 breast carcinoma cells; 'ATCC HTB22'.

PN WO8907107-A.

PD 10-AUG-1989.

PF 29-JAN-1988; 88US-00149831.

PR 29-JAN-1988; 88US-00149831.

PA (DANA-) DANA-FARBER CANCER.

PI Kufe DW;

DR WPI; 1989-248989/34.

DR P-PSDB; AAP91045, AAP91053, AAP91054, AAP91046.

PT Recombinant polypeptide(s) - contains DF 3 breast carcinoma antigen

PT epitope and useful as assay reagents, and encoding DNA sequences.

PS Claim 1; Fig 4; 31pp; English.

CC The sequence encodes a carbohydrate-free polypeptide contg. DF3 breast
 CC carcinoma antigen epitope. Useful as a competitive binding assay reagent
 CC and improves diagnosis. The cDNA consists of nearly identical 60 BP
 CC tandem repeats which are 85% GC-rich. See also AAP91045, AAP91046,
 CC AAP91053, AAP91054. (Updated on 25-MAR-2003 to correct PA field.)
 CC (Updated on 27-AUG-2003 to correct OS field.)

Sequence 309 BP; 21 A; 92 C; 160 G; 36 T; 0 U; 0 Other;

Query Match 100.0%; Score 60; DB 1; Length 309;
 Best Local Similarity 100.0%; Pred. No. 1.6e-06;
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCTCCACCGCCCGCCCGCCGCGGTGTACCTCGGCGCCGACACCGAGCGCGCCCG 60
 DB 266 GGCTCCACCGCCCGCCCGCCGCGGTGTACCTCGGCGCCGACACCGAGCGCGCCCG 207

RESULT 3

AD157712
 ID AD157712 standard; cDNA; 1194 BP.

AC AD157712;

DT 22-APR-2004 (first entry)

DE Human breast specific nucleic acid (BSNA) #83.

KW Human; breast specific nucleic acid; BSNA; gene; ss; metastasis;

OS Homo sapiens.

PN WO2003106648-A2.

PD 24-DEC-2003.

PF 16-JUN-2003; 2003WO-US018934.

PR 14-JUN-2002; 2002US-0389327P.

PA (DIAD-) DIADEXUS INC.

PI Salceda S, Macina RA, Turner LR, Sun Y, Liu C;

DR WPI; 2004-082185/08.

DR P-PSDB; AD157782.

PT Novel isolated polypeptide comprising breast specific protein sequences,
 PT useful for diagnosing or monitoring presence and metastases of breast
 PT cancer in patient.

PS Claim 1; SEQ ID NO 83; 370pp; English.

CC The invention relates to human breast specific nucleic acids (BSNA) and
 CC the breast specific proteins (BSP) they encode. The nucleic acids are
 CC useful for determining the presence of a BSNA in a sample which involves
 CC selecting the sample with a BSNA under conditions in which the BSNA will
 CC selectively hybridise to a BSNA in the sample, and detecting the
 CC of a BSP in a sample which involves determining the presence
 CC of a reagent under conditions in which the reagent will selectively interact
 CC with the BSP, and detecting the interaction of the reagent with a BSP in
 CC the sample. The nucleic acids and proteins are useful for diagnosing or
 CC monitoring the presence and metastases of breast cancer in a patient,
 CC which involves determining an amount of nucleic acid or protein and
 CC comparing the determined amount of nucleic acid or protein in the sample
 CC of the patient to the amount of a breast specific marker in a normal
 CC control, where a difference in the determined amount in the sample
 CC compared to the amount in the control is associated with the presence of
 CC breast cancer. The sequences are useful for treating a patient with
 CC a BSP to a patient, where the administration induces an immune
 CC response against the breast cancer cell expressing the BSNA or BSP.
 CC Sequence represents a human BSNA of the invention.

Sequence 1194 BP; 243 A; 438 C; 275 G; 238 T; 0 U; 0 Other;

Query Match 100.0%; Score 60; DB 12; Length 1194;
 Best Local Similarity 100.0%; Pred. No. 1.4e-06;
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCTCCACCGCCCGCCCGCCGCGGTGTACCTCGGCGCCGACACCGAGCGCGCCCG 60
 DB 556 GGCTCCACCGCCCGCCCGCCGCGGTGTACCTCGGCGCCGACACCGAGCGCGCCCG 655

RESULT 4

AD157693
 ID AD157693 standard; cDNA; 1378 BP.

AC AD157693;

XX 22-APR-2004 (first entry)
 XX Human breast specific nucleic acid (BSNA) #64.
 DE Human breast specific nucleic acid; BSNA; gene; ss; metastasis;
 XX breast cancer; cytostatic.
 XX Homo sapiens.
 OS
 XX WO2003106648-A2.
 PN
 XX 24-DEC-2003.
 PD
 XX 16-JUN-2003; 2003WO-US018934.
 PF
 XX 14-JUN-2002; 2002US-0389327P.
 PR
 XX (DIAD-) DIADEXUS INC.
 PA
 XX Salceda S, Macina RA, Turner LR, Sun Y, Liu C;
 PI
 XX WPI; 2004-082185/08.
 DR
 XX P-PSDB; AD157765.
 DR
 XX Novel isolated polypeptide comprising breast specific protein sequences,
 PT useful for diagnosing or monitoring presence and metastases of breast
 PT cancer in patient.
 PS
 XX Claim 1; SEQ ID NO 64; 370pp; English.
 PS
 XX The invention relates to human breast specific nucleic acids (BSNA) and
 CC the breast specific proteins (BSP) they encode. The nucleic acids are
 CC useful for determining the presence of a BSNA in a sample which involves
 CC contacting the sample with a BSNA under conditions in which the BSNA will
 CC selectively hybridize to a BSNA in the sample, and detecting the
 CC hybridization. The nucleic acids are useful for determining the presence
 CC of a BSP in a sample which involves contacting the sample with suitable
 CC reagent under conditions in which the reagent will selectively interact
 CC with the BSP, and detecting the interaction of the reagent with a BSP in
 CC the sample. The nucleic acids and proteins are useful for diagnosing or
 CC monitoring the presence and metastases of breast cancer in a patient,
 CC which involves determining an amount of nucleic acid or protein and
 CC comparing the determined amount of nucleic acid or protein in the sample
 CC of the patient to the amount of a breast specific marker in a normal
 CC control, where a difference in the determined amount in the sample
 CC compared to the amount in the control is associated with the presence of
 CC breast cancer. The sequences are useful for treating a patient with
 CC breast cancer, involving administering a composition consisting of a BSNA
 CC or a BSP to a patient, where the administration induces an immune
 CC response against the breast cancer cell expressing the BSNA or BSP. This
 CC sequence represents a human BSNA of the invention.
 CC
 SQ Sequence 1378 BP; 295 A; 489 C; 316 G; 278 T; 0 U; 0 Other;
 SQ
 Query Match 100.0%; Score 60; DB 12; Length 1378;
 Best Local Similarity 100.0%; Pred. No. 1.4e-06;
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGCTCACCAGCCCCCGCCAGCCGAGTGTACCTCGGCCCCGGAGACACGAGCGGCCCG 60
 DB 596 GGCTCACCAGCCCCCGCCAGCCGAGTGTACCTCGGCCCCGGAGACACGAGCGGCCCG 655
 RESULT 5
 ID ADO23180/c
 XX ADO23180 standard; RNA; 1424 BP.
 AC ADO23180;
 XX
 XX 12-AUG-2004 (first entry)
 DT
 XX Antisense human MUC1 mucin glycoprotein RNA (coding sequence) Seqid 75.
 DE

XX human; MUC1; mucin glycoprotein; cancer; chemotherapeutic; MUC1/BCD; ss;
 KM antisense.
 KW
 XX Homo sapiens.
 OS
 XX WO2004044160-A2.
 PN
 XX 27-MAY-2004.
 PD
 XX 12-NOV-2003; 2003WO-US035848.
 PF
 XX 13-NOV-2002; 2002US-00293391.
 PR
 XX 29-MAY-2003; 2003US-00447839.
 PR
 XX (DAND) DANA FARBER CANCER INST INC.
 PA
 XX (ILEX-) ILEX PROD INC.
 PI
 XX Kufe DW, Kharbanda S, Weltman SD;
 PI
 XX WPI; 2004-420304/39.
 DR
 XX Double-stranded RNA complex useful for inhibiting proliferation of cancer
 PT cell expressing MUC1 mucin glycoprotein, comprises first and second RNA
 PT sequences.
 PS
 XX Disclosure; SEQ ID NO 75; 112pp; English.
 PS
 XX This invention relates to novel modulators of the human MUC1 mucin
 CC glycoprotein for use in cancer therapeutics, where MUC1 is a protein that
 CC acts to inhibit the apoptotic response to genotoxic stress caused by
 CC chemotherapeutic agents. In particular, it refers to modulators of the
 CC MUC1 extracellular domain (MUC1/BCD). The method refers to using double-
 CC stranded RNA complexes as MUC1 interference RNA compositions such that
 CC MUC1 expression is inhibited, which in turn inhibits cancer cell
 CC proliferation. The present invention describes screening assays to
 CC identify compounds that inhibit the binding of various MUC1 ligands such
 CC as neutroglin 2, as well as agonists, antagonists and antibodies thereof.
 CC Furthermore, it provides MUC1 antisense, small interfering RNA (siRNA)
 CC and small molecules in combination with chemotherapeutic agents that are
 CC useful in the field of cancer therapy. This polynucleotide sequence is
 CC the antisense human MUC1 RNA of the invention.
 CC
 SQ Sequence 1424 BP; 309 A; 325 C; 495 G; 0 T; 295 U; 0 Other;
 SQ
 Query Match 100.0%; Score 60; DB 12; Length 1424;
 Best Local Similarity 100.0%; Pred. No. 1.4e-06;
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGCTCACCAGCCCCCGCCAGCCGAGTGTACCTCGGCCCCGGAGACACGAGCGGCCCG 60
 DB 1041 GGCTCACCAGCCCCCGCCAGCCGAGTGTACCTCGGCCCCGGAGACACGAGCGGCCCG 982
 RESULT 6
 ID ABL60159
 XX ABL60159 standard; cDNA; 1428 BP.
 AC ABL60159;
 XX
 XX 22-JUL-2002 (first entry)
 DT
 XX Human MUC1 encoding cDNA SEQ ID NO 2.
 DE
 XX Human; mucin 1; MUC1; transmembrane protein; SNP; cancer; cytostatic;
 KW single nucleotide polymorphism; haplotyping; genotyping; drug;
 KW antiinflammatory; gene; ss.
 XX
 XX Homo sapiens.
 OS
 XX
 XX Key Location/Qualifiers
 FH CDS 1..1428
 FT
 FT /*tag= a

```

FT      /product= "MUC1"
FT      replace(1009,A)
FT      /*tag= b
FT      /standard name= "Single nucleotide polymorphism"
FT      /note= "SNP allelic variation results in Val substituted
FT      by Met at position 337 of the MUC1 protein (ABB77476)"
XX      WO200226765-A2.
XX
XX      04-APR-2002.
XX
XX      25-SEP-2001; 2001WO-US030151.
XX
XX      28-SEP-2000; 2000US-0236113P.
XX
XX      (GENA-) GENAISSANCE PHARM INC.
XX
XX      Chew A, Koshiy B;
XX
XX      WPI, 2002-405042/43.
XX      P-PSDB; ABB77476.
XX
XX      New genetic variants of mucin 1, Transmembrane gene, useful in studying
XX      expression and function of protein encoded by the gene and for screening
XX      drugs to treat diseases e.g. cancer.
XX
XX      Claim 23; Fig 2; 75pp; English.
XX
XX      The invention relates to a polynucleotide (ABU60158, ABU60159) encoding
XX      mucin 1/MUC1 (ABB77476), Transmembrane isogene. The invention describes
XX      novel genetic variants of the MUC1 gene. The invention is useful for
XX      haployping/genotyping the MUC1 gene in an individual and identifying an
XX      association between a trait and at least one of the haplotypes or
XX      haploype pairs of MUC1 gene. MUC1 is useful for studying the expression
XX      and function of MUC1 and expressing MUC1 protein for use in screening for
XX      candidate drugs to treat diseases related to MUC1 activity and in
XX      studying the effect of the variation on the biological activity of MUC1
XX      as well as on the binding affinity of candidate drugs targeting MUC1 for
XX      the treatment of e.g. cancer. MUC1 is further used by the pharmaceutical
XX      research scientist to validate MUC1 as a candidate target for and in
XX      design of clinical trials of candidate drugs for treating a specific
XX      condition drugs or disease predicted to be associated with MUC1 activity.
XX      MUC1 antibodies are useful in a variety of diagnostic and prognostic
XX      formats and therapeutic methods
XX
XX      Sequence 1428 BP; 296 A; 496 C; 327 G; 309 T; 0 U; 0 Other;
XX
XX      Query Match      100.0%; Score 60; DB 6; Length 1428;
XX      Best Local Similarity 100.0%; Pred. No. 1,4e-06;
XX      Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX      QY      1 GGCTCCACCGCCCCCGAGCCACGGTGTCACTCGGCCCGGACACGAGCGGCCCG 60
XX      DB      385 GGCTCCACCGCCCCCGAGCCACGGTGTCACTCGGCCCGGACACGAGCGGCCCG 444
XX
XX      RESULT 7
XX      ADO33125
XX      ID      ADO23125 standard; RNA; 1428 BP.
XX
XX      AC      ADO23125;
XX
XX      DT      12-AUG-2004 (first entry)
XX
XX      DE      Human MUC1 mucin glycoprotein RNA (coding sequence) SeqID 20.
XX
XX      KW      human; MUC1; mucin glycoprotein; cancer; chemotherapeutic; MUC1/ECG; ss.
XX
XX      OS      Homo sapiens.
XX
XX      PN      WO2004044160-A2.
XX
XX      PD      27-MAY-2004.

```

```

XX      12-NOV-2003; 2003WO-US035848.
XX
XX      13-NOV-2002; 2002US-00293391.
XX      29-MAY-2003; 2003US-00447839.
XX
XX      (DAND ) DANA FARBER CANCER INST INC.
XX      (ILEX-) ILEX PROD INC.
XX
XX      Kufe DW, Kharbada S, Weltman SD;
XX
XX      WPI; 2004-420304/39.
XX
XX      Double-stranded RNA complex useful for inhibiting proliferation of cancer
XX      cell expressing MUC1 mucin glycoprotein, comprises first and second RNA
XX      sequences.
XX
XX      Claim 2; SEQ ID NO 20; 112pp; English.
XX
XX      This invention relates to novel modulators of the human MUC1 mucin
XX      glycoprotein for use in cancer therapeutics, where MUC1 is a protein that
XX      acts to inhibit the apoptotic response to genotoxic stress caused by
XX      chemotherapeutic agents. In particular, it refers to modulators of the
XX      MUC1 extracellular domain (MUC1/ECG). The method refers to using double-
XX      stranded RNA complexes as MUC1 interference RNA compositions such that
XX      MUC1 expression is inhibited, which in turn inhibits cancer cell
XX      proliferation. The present invention describes screening assays to
XX      identify compounds that inhibit the binding of various MUC1 ligands such
XX      as neutregulin 2, as well as agonists, antagonists and antibodies thereof.
XX      Furthermore, it provides MUC1 antisense, small interfering RNA (siRNA)
XX      and small molecules in combination with chemotherapeutic agents that are
XX      useful in the field of cancer therapy. This polynucleotide sequence is
XX      the human MUC1 RNA of the invention.
XX
XX      Sequence 1428 BP; 296 A; 496 C; 327 G; 0 T; 309 U; 0 Other;
XX
XX      Query Match      100.0%; Score 60; DB 12; Length 1428;
XX      Best Local Similarity 93.3%; Pred. No. 1,4e-06;
XX      Matches 56; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
XX
XX      QY      1 GGCTCCACCGCCCCCGAGCCACGGTGTCACTCGGCCCGGACACGAGCGGCCCG 60
XX      DB      385 GGCTCCACCGCCCCCGAGCCACGGTGTCACTCGGCCCGGACACGAGCGGCCCG 444
XX
XX      RESULT 8
XX      ADF32627
XX      ID      ADF32627 standard; DNA; 1457 BP.
XX
XX      AC      ADF32627;
XX
XX      DT      26-FEB-2004 (first entry)
XX
XX      DE      Plasmid JNW358 MUC-1 nucleotide sequence.
XX
XX      MUC-1 antigen; immune response; MUC-1; variable number of tandem repeat;
XX      VNTR; repeat unit; tumour; metastasis; cyrostatic; vaccine; gene therapy;
XX      gene; ds.
XX
XX      OS      Synthetic.
XX
XX      PN      WO2003100060-A2.
XX
XX      PD      04-DEC-2003.
XX
XX      23-MAY-2003; 2003WO-EP005594.
XX      24-MAY-2002; 2002GB-00012046.
XX      (GLAX ) GLAXO GROUP LTD.
XX
XX      Burden N, Ellis JH, Hamblin PA;

```

DR WPI; 2004-042811/04.
XX
XX New nucleic acid molecule encoding a MUC-1 antigen, useful for preparing
PT a composition for treating or preventing tumors or metastases.
XX
XX Example; Fig 3; 66pp; English.
XX
CC The present invention describes a nucleic acid molecule which encodes a
CC MUC-1 antigen. The nucleic acid is capable of raising an immune response
CC in vivo, has reduced susceptibility to recombination than full-length MUC
CC -1 and comprises between 1 and 15 variable number of tandem repeats
CC (VNR) perfect repeat units. Also described: (1) a plasmid comprising the
CC DNA molecule; (2) a protein encoded by the nucleic acid; (3) a
CC pharmaceutical composition comprising the nucleic acid, plasmid or
CC protein and an excipient, diluent or carrier; and (4) a method of
CC treating or preventing tumors or metastases. A MUC1 antigen has
CC cytostatic activity, and can be used in vaccines, and in gene therapy.
CC The nucleic acid is useful for preparing a composition for treating or
CC preventing tumors or metastases. The present sequence is used in the
CC exemplification of the present invention.
XX
SQ Sequence 1457 BP; 302 A; 506 C; 334 G; 315 T; 0 U; 0 Other;
XX
Query Match 100.0%; Score 60; DB 12; Length 1457;
Best Local Similarity 100.0%; Pred. No. 1.4e-06;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
OY 1 GGCTCCACCGCCCCCGCCAGCCGCGTGTACCTCGGCCCGAGACACAGCGGCCCGG 60
DB 403 GGCTCCACCGCCCCCGCCAGCCGCGTGTACCTCGGCCCGAGACACAGCGGCCCGG 462
XX
RESULT 9
AAV48329
ID AAV48329 standard; cDNA; 1527 BP.
XX
XX AAV48329;
XX
XX 20-NOV-1998 (first entry)
XX
XX MiniMUC1 gene.
XX
XX 89; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;
XX tumour; tumour-associated antigen.
XX
XX Homo sapiens.
XX
XX OS
XX FH
XX FT
XX CDS
XX
XX Key Location/Qualifiers
XX 1.1527
XX /*tag= a
XX /product= "MiniMUC1 protein"
XX
XX MO9837095-A2.
XX
XX 27-AUG-1998.
XX
XX 24-FEB-1998; 98WO-US003693.
XX
XX 24-FEB-1997; 97US-0038253P.
XX
XX (THER-) THERION BIOLOGICS CORP.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX (DAND) DANA FARBER CANCER INST INC.
XX
XX Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;
XX WPI; 1998-467492/40.
XX P-PSDB; AAN77233.
XX
XX New recombinant pox virus for tumour therapy - comprises DNA encoding an
XX immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
XX
XX Example 1; Page 21-22; 42pp; English.

XX
XX The immunogenic mini-MUC1 fragment was created from MUC1 tandem repeat
CC units for inclusion in a recombinant pox virus (RPV). The RPV was used in
CC a pharmaceutical composition also containing an immunomodulator to
CC generate MUC1 specific cytotoxic T-lymphocytes. The recombinant pox virus
CC therefore encodes an immunogenic MUC1 fragment that does not undergo
CC significant genetic deletion, thereby providing an unexpectedly stable
CC and immunogenic pox virus. They can be used to prevent or treat tumours
CC expressing MUC1 tumour-associated antigens
XX
SQ Sequence 1527 BP; 296 A; 573 C; 351 G; 307 T; 0 U; 0 Other;
XX
Query Match 100.0%; Score 60; DB 2; Length 1527;
Best Local Similarity 100.0%; Pred. No. 1.4e-06;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
OY 1 GGCTCCACCGCCCCCGCCAGCCGCGTGTACCTCGGCCCGAGACACAGCGGCCCGG 60
DB 226 GGCTCCACCGCCCCCGCCAGCCGCGTGTACCTCGGCCCGAGACACAGCGGCCCGG 285
XX
RESULT 10
ADK70370
ID ADK70370 standard; cDNA; 1614 BP.
XX
XX ADK70370;
XX
XX 06-MAY-2004 (first entry)
XX
XX Respiratory disease differentially expressed cDNA #106.
XX
XX ds; gene; cytostatic; respiratory; antiasthmatic; gene therapy;
XX differential gene expression; respiratory disorder; lung cancer;
XX chronic obstructive pulmonary disease; emphysema; asthma.
XX
XX Homo sapiens.
XX
XX OS
XX FH
XX FT
XX CDS
XX
XX Key Location/Qualifiers
XX 1.1527
XX /*tag= a
XX /product= "MiniMUC1 protein"
XX
XX MO9837095-A2.
XX
XX 27-AUG-1998.
XX
XX 24-FEB-1998; 98WO-US003693.
XX
XX 24-FEB-1997; 97US-0038253P.
XX
XX (THER-) THERION BIOLOGICS CORP.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX (DAND) DANA FARBER CANCER INST INC.
XX
XX Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;
XX WPI; 1998-467492/40.
XX P-PSDB; AAN77233.
XX
XX New recombinant pox virus for tumour therapy - comprises DNA encoding an
XX immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.
XX
XX Example 1; Page 21-22; 42pp; English.

ID ADI57686 standard; cDNA; 1712 BP.
 XX
 AC ADI57686;
 XX
 DT 22-APR-2004 (first entry)
 XX
 DE Human breast specific nucleic acid (BSNA) #57.
 XX
 KM Human; breast specific nucleic acid; BSNA; gene; ss; metastasis;
 KM breast cancer; cytosstatic.
 XX
 OS Homo sapiens.
 XX
 PN MO2003106648-A2.
 XX
 PD 24-DEC-2003.
 XX
 PF 16-JUN-2003; 2003WO-US018934.
 XX
 PR 14-JUN-2002; 2002US-0389327P.
 XX
 PA (DIAD-) DIADEXUS INC.
 XX
 PI Salceda S, Macina RA, Turner LR, Sun Y, Liu C;
 XX
 DR MPI; 2004-082185/08.
 DR P-PSDB; ADI57758.
 XX
 PT Novel isolated polypeptide comprising breast specific protein sequences,
 PT useful for diagnosing or monitoring presence and metastases of breast
 PT cancer in patient.
 XX
 PS Claim 1; SEQ ID NO 57; 370pp; English.
 XX
 CC The invention relates to human breast specific nucleic acids (BSNA) and
 CC the breast specific proteins (BSP) they encode. The nucleic acids are
 CC useful for determining the presence of a BSNA in a sample which involves
 CC contacting the sample with a BSNA under conditions in which the BSNA will
 CC selectively hybridize to a BSNA in the sample, and detecting the presence
 CC of a BSP in a sample which involves contacting the sample with suitable
 CC reagent under conditions in which the reagent will selectively interact
 CC with the BSP, and detecting the interaction of the reagent with a BSP in
 CC the sample. The nucleic acids and proteins are useful for diagnosing or
 CC monitoring the presence and metastases of breast cancer in a patient,
 CC which involves determining an amount of nucleic acid or protein and
 CC comparing the determined amount of nucleic acid or protein in the sample
 CC of the patient to the amount of a breast specific marker in a normal
 CC control, where a difference in the determined amount in the sample
 CC compared to the amount in the control is associated with the presence of
 CC breast cancer. The sequences are useful for treating a patient with
 CC breast cancer, involving administering a composition consisting of a BSNA
 CC or a BSP to a patient, where the administration induces an immune
 CC response against the breast cancer cell expressing the BSNA or BSP. This
 CC sequence represents a human BSNA of the invention.
 XX
 SQ Sequence 1712 BP; 327 A; 604 C; 435 G; 346 T; 0 U; 0 Other;
 XX
 Query Match 100.0%; Score 60; DB 12; Length 1712;
 Best Local Similarity 100.0%; Pred. No. 1.4e-06;
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGCTCCACCGCCCGCCGAGGTGTCACCTCGGCCCCGGACACGAGCGGCCCG 60
 DB 596 GGCTCCACCGCCCGCCGAGGTGTCACCTCGGCCCCGGACACGAGCGGCCCG 655
 XX
 AC ADI57669
 ID ADI57669 standard; cDNA; 1738 BP.
 XX
 AC ADI57669;
 XX

DT 22-APR-2004 (first entry)
 XX
 DE Human breast specific nucleic acid (BSNA) #40.
 XX
 KM Human; breast specific nucleic acid; BSNA; gene; ss; metastasis;
 KM breast cancer; cytosstatic.
 XX
 OS Homo sapiens.
 XX
 PN MO2003106648-A2.
 XX
 PD 24-DEC-2003.
 XX
 PF 16-JUN-2003; 2003WO-US018934.
 XX
 PR 14-JUN-2002; 2002US-0389327P.
 XX
 PA (DIAD-) DIADEXUS INC.
 XX
 PI Salceda S, Macina RA, Turner LR, Sun Y, Liu C;
 XX
 DR MPI; 2004-082185/08.
 DR P-PSDB; ADI57743.
 XX
 PT Novel isolated polypeptide comprising breast specific protein sequences,
 PT useful for diagnosing or monitoring presence and metastases of breast
 PT cancer in patient.
 XX
 PS Claim 1; SEQ ID NO 40; 370pp; English.
 XX
 CC The invention relates to human breast specific nucleic acids (BSNA) and
 CC the breast specific proteins (BSP) they encode. The nucleic acids are
 CC useful for determining the presence of a BSNA in a sample which involves
 CC contacting the sample with a BSNA under conditions in which the BSNA will
 CC selectively hybridize to a BSNA in the sample, and detecting the
 CC hybridization. The nucleic acids are useful for determining the presence
 CC of a BSP in a sample which involves contacting the sample with suitable
 CC reagent under conditions in which the reagent will selectively interact
 CC with the BSP, and detecting the interaction of the reagent with a BSP in
 CC the sample. The nucleic acids and proteins are useful for diagnosing or
 CC monitoring the presence and metastases of breast cancer in a patient,
 CC which involves determining an amount of nucleic acid or protein and
 CC comparing the determined amount of nucleic acid or protein in the sample
 CC of the patient to the amount of a breast specific marker in a normal
 CC control, where a difference in the determined amount in the sample
 CC compared to the amount in the control is associated with the presence of
 CC breast cancer. The sequences are useful for treating a patient with
 CC breast cancer, involving administering a composition consisting of a BSNA
 CC or a BSP to a patient, where the administration induces an immune
 CC response against the breast cancer cell expressing the BSNA or BSP. This
 CC sequence represents a human BSNA of the invention.
 XX
 SQ Sequence 1738 BP; 342 A; 642 C; 399 G; 355 T; 0 U; 0 Other;
 XX
 Query Match 100.0%; Score 60; DB 12; Length 1738;
 Best Local Similarity 100.0%; Pred. No. 1.4e-06;
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGCTCCACCGCCCGCCGAGGTGTCACCTCGGCCCCGGACACGAGCGGCCCG 60
 DB 458 GGCTCCACCGCCCGCCGAGGTGTCACCTCGGCCCCGGACACGAGCGGCCCG 517
 XX
 AC ADI57673;
 ID ADI57673 standard; cDNA; 1755 BP.
 XX
 AC ADI57673;
 XX
 DT 22-APR-2004 (first entry)
 XX
 DE Human breast specific nucleic acid (BSNA) #44.
 XX

KW Human: breast specific nucleic acid; BSNA; gene; ss; metastasis;
 KW breast cancer; cytostatic.
 XX
 OS Homo sapiens.
 XX
 PN WO2003106648-A2.
 XX
 PD 24-DEC-2003.
 XX
 PF 16-JUN-2003; 2003WO-US018934.
 XX
 PR 14-JUN-2002; 2002US-0389327P.
 XX
 PA (DIAD-) DIADEXUS INC.
 PI Salceda S, Macina RA, Turner LR, Sun Y, Liu C;
 XX
 DR MPI; 2004-082185/08.
 DR P-PSDB; ADI5746.
 XX
 PT Novel isolated polypeptide comprising breast specific protein sequences,
 PT useful for diagnosing or monitoring presence and metastases of breast
 PT cancer in patient.
 XX
 PS Claim 1; SEQ ID NO 44; 370pp; English.
 XX
 CC The invention relates to human breast specific nucleic acids (BSNA) and
 CC the breast specific proteins (BSP) they encode. The nucleic acids are
 CC useful for determining the presence of a BSNA in a sample which involves
 CC contacting the sample with a BSNA under conditions in which the BSNA will
 CC selectively hybridise to a BSNA in the sample, and detecting the
 CC hybridisation. The nucleic acids are useful for determining the presence
 CC of a BSP in a sample which involves contacting the sample with suitable
 CC reagent under conditions in which the reagent will selectively interact
 CC with the BSP, and detecting the interaction of the reagent with a BSP in
 CC the sample. The nucleic acids and proteins are useful for diagnosing or
 CC monitoring the presence and metastases of breast cancer in a patient,
 CC which involves determining an amount of nucleic acid or protein and
 CC comparing the determined amount of nucleic acid or protein in the sample
 CC of the patient to the amount of a breast specific marker in a normal
 CC control, where a difference in the determined amount in the sample
 CC compared to the amount in the control is associated with the presence of
 CC breast cancer. The sequences are useful for treating a patient with
 CC breast cancer, involving administering a composition consisting of a BSNA
 CC or a BSP to a patient, where the administration induces an immune
 CC response against the breast cancer cell expressing the BSNA or BSP. This
 CC sequence represents a human BSNA of the invention.
 XX
 SQ Sequence 1755 BP; 348 A; 595 C; 432 G; 378 T; 0 U; 2 Other;

Query Match 100.0%; Score 60; DB 12; Length 1755;
 Best Local Similarity 100.0%; Pred. No. 1.4e-06;
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCTCCACCGCCCCCGACGACGATGTCACCTCGGCCCGGACACACGCGGCCCG 60
 |||
 DB 569 GGCTCCACCGCCCCCGACGACGATGTCACCTCGGCCCGGACACACGCGGCCCG 628
 |||

Search completed: January 15, 2005, 20:36:05
 Job time : 171.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: January 15, 2005, 19:43:55 ; Search time 1550.6 Seconds
(without alignments)
1410.024 Million cell updates/sec

Title: US-10-057-136-2

Perfect score: 60

Sequence: 1 ggcacacagcccccacccagc.....cgacacacagccgcccgcg 60

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 10%

Listing first 45 summaries

Database :

EST: *
1: gb_esc1: *
2: gb_esc2: *
3: gb_hic: *
4: gb_esc3: *
5: gb_esc4: *
6: gb_esc5: *
7: gb_esc6: *
8: gb_gss1: *
9: gb_gss2: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	330	1	AI925867 wo20d04.x
2	60	100.0	604	4	BM791359 K-EST0071
3	60	100.0	877	5	BU542454 AGENCOURT
4	60	100.0	959	6	CA489836 AGENCOURT
5	60	100.0	1113	6	BU148487 AGENCOURT
6	60	100.0	1130	5	BU542996 AGENCOURT
7	60	100.0	1234	5	BQ936898 AGENCOURT
8	60	100.0	1268	5	BQ943554 AGENCOURT
9	60	100.0	1343	5	BQ920055 AGENCOURT
10	60	100.0	1349	5	BU152566 AGENCOURT
11	60	100.0	1420	5	BU542790 AGENCOURT
12	58.4	97.3	1531	5	BU543309 AGENCOURT
13	56.8	94.7	1531	5	BQ935496 AGENCOURT
14	52.2	87.0	1536	5	BQ923149 AGENCOURT
15	52.2	87.0	1536	5	BQ923149 AGENCOURT
16	51.2	85.3	166	7	T27692
17	51.2	85.3	475	6	CB120860 K-EST0168
18	51.2	85.3	669	6	CB122585 K-EST0170
19	50.4	84.0	1578	4	BG775565 602650481
20	49.4	82.3	1334	5	BQ943809 AGENCOURT
21	48.8	81.3	472	4	BM759495 K-EST0039
22	46	76.7	619	4	BI260921 602970962
23	34.2	57.0	412	5	BQ082925 K-EST0144
24	34.2	57.0	670	2	BF338440 602034094

25	34.2	57.0	690	1	AL543598
26	34.2	57.0	1241	3	CR596859 full-1-1eng
27	32.8	54.7	548	2	BE706360 RCI-HT025
28	32.6	54.3	400	6	CB667180 OSJNRd14C
29	32.6	54.3	663	6	CB667142 OSJNRd14F
30	32.6	53.3	572	6	CA631630 wle1n.pk0
31	32	53.3	727	2	BF203237 601866065
32	31.8	53.0	308	7	R48730 Y168a12.r1
33	31.2	52.0	1196	5	BQ050961 AGENCOURT
34	31	51.7	905	2	BE216819 HY_CEB001
35	30.8	51.3	1306	4	BM909484 AGENCOURT
36	30.8	51.3	1716	9	CL507815 SA1L_788
37	30.4	50.7	1004	8	B13383 jdl51_Typp
38	30.4	50.7	1178	8	CC227732 CH261-156
39	30.2	50.3	947	6	CB200158 AGENCOURT
40	30.2	50.3	1080	8	AQ893056 HS_4832_A
41	30.2	50.3	1436	4	B1143552 AU182461
42	30	50.0	240	1	BU756102 SALX_0530
43	29.8	49.7	363	4	BI132432 AR030D06L
44	29.8	49.7	437	8	BI132432 AR030D06L
45	29.8	49.7	471	1	AJ683668 AJ683668

ALIGNMENTS

RESULT 1
AI925867/c 330 bp mRNA linear EST 08-MAR-2000
LOCUS wo20d04.x1 NCI CGAP Pauli Homo sapiens cDNA clone IMAGE:2455879 3'
DEFINITION similar to contig element MSRI MSRI repetitive element ; mRNA

ACCESSION AI925867 GI:5661831
VERSION AI925867.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE

1 (bases 1 to 330)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Life Technologies catalog #: 11548-013
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.lnl.gov/bbrp/image/image.html
Insert Length: 578 Std Error: 0.00
Seq primer: -40up from Glibco
High quality sequence stop: 305.
Location/Qualifiers
1..330
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2455879"
/tissue_type="adenocarcinoma"
/lab_host="DH10B"
/clone_lib="NCI-CGAP_Pauli"
/note="Organ: pancreas; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.72 kb. Life Technologies catalog #: 11548-013"

JOURNAL

COMMENT

FEATURES

source

ORIGIN

Query Match 100.0%; Score 60; DB 1; Length 330;
Best Local Similarity 100.0%; Pred. No. 6.3e-06;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGCTCCACGCGCCCCCAGCCCGGTGTCACTCGGCGCCCGGACACCAAGCGCCCGC 60
 |||||
 DB 316 GGCTCCACGCGCCCCCAGCCCGGTGTCACTCGGCGCCCGGACACCAAGCGCCCGC 257
 |||||
 RESULT 2
 LOCUS BM791359 604 bp mRNA linear EST 05-MAR-2002
 DEFINITION K-EST0071342 S21SNUS20 Homo sapiens cDNA clone S21SNUS20-14-A06 5',
 mRNA sequence.
 ACCESSION BM791359
 VERSION BM791359.1 GI:19139591
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 604)
 Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
 Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
 Kim,Y.S.
 21C Frontier Korean EST Project 2001
 Unpublished (2002)
 CONTACT Contact: Kim YS
 Genome Research Center
 Korea Research Institute of Bioscience & Biotechnology
 52 Eosun-dong Yuseong-gu, Daejeon 305-333, South Korea
 Tel: +82-42-860-4470
 Fax: +82-42-860-4409
 Email: yongsung@mail.krrib.re.kr
 Plate: 14 row: A column: 06
 High quality sequence stop: 604.
 Location/Qualifiers
 1..604
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="S21SNUS20-14-A06"
 /sex="F"
 /tissue_type="Stomach"
 /cell_type="floating aggregates"
 /cell_line="SNU-520"
 /lab_host="Top10F"
 /clone_lib="S21SNUS20"
 /note="Organ: Stomach; Vector: pTZ18RP1; Site 1: EcoRI;
 Site 2: NotI; The poly (A)+ RNA was dephosphorylated with
 bacterial alkaline phosphatase (BAP) and then deapped
 with tobacco acid pyrophosphatase (PAP). The deapped
 intact mRNA was ligated with DNA-RNA linker including EcoR
 I site by treatment of T4 RNA ligase and the first strand
 cDNA was synthesized from oligo dT-selected mRNA by
 priming with dT-tailed vector. The dT-tailed vector was
 adjusted to have about 60nt. The cDNA vector was
 circularized with E. coli DNA ligase after digestion of
 EcoRI which site is also included in vector. An RNA strand
 converted to a DNA strand by Okayama-Berg method. The
 obtained cDNA vectors were used for transformation of
 competent cells E. coli Top10F' by electroporation method.
 The cDNA libraries constructed by this method are
 full-length enriched cDNA library."

Query Match 100.0%; Score 60; DB 4; Length 604;
 Best Local Similarity 100.0%; Pred. No. 5.9e-06;
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGCTCCACGCGCCCCCAGCCCGGTGTCACTCGGCGCCCGGACACCAAGCGCCCGC 60
 |||||
 DB 106 GGCTCCACGCGCCCCCAGCCCGGTGTCACTCGGCGCCCGGACACCAAGCGCCCGC 165
 |||||

RESULT 3
 BUS42454

LOCUS BUS42454 877 bp mRNA linear EST 13-SEP-2002
 DEFINITION AGENCOURT_10322173 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6574322
 5', mRNA Sequence.
 ACCESSION BUS42454
 VERSION BUS42454.1 GI:22852937
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 877)
 NIH-MGC http://mgs.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-r@mail.nih.gov
 Tissue Procurement: DCTD/DRP
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LINC2769 row: h column: 02
 High quality sequence stop: 760.
 Location/Qualifiers
 1..877
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6574322"
 /tissue_type="carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_40"
 /note="Organ: prostate; Vector: pOTB7; Site 1: XhoI;
 Site 2: EcoRI; cDNA made by oligo-dT priming.
 directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCAAGAG(G). Library constructed by
 Ling Hong in the Laboratory of Gerald W. Rubin (University
 of California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library."

Query Match 100.0%; Score 60; DB 5; Length 877;
 Best Local Similarity 100.0%; Pred. No. 5.7e-06;
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGCTCCACGCGCCCCCAGCCCGGTGTCACTCGGCGCCCGGACACCAAGCGCCCGC 60
 |||||
 DB 35 GGCTCCACGCGCCCCCAGCCCGGTGTCACTCGGCGCCCGGACACCAAGCGCCCGC 94
 |||||

RESULT 4
 LOCUS CA489836 959 bp mRNA linear EST 14-NOV-2002
 DEFINITION AGENCOURT_10810668 MAPCL Homo sapiens cDNA clone IMAGE:6722324 5',
 mRNA sequence.
 ACCESSION CA489836
 VERSION CA489836.1 GI:24952627
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 959)
 NIH-MGC http://mgs.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-r@mail.nih.gov
 Tissue Procurement: Kristi A. Eglund, Ira Pastan
 cDNA Library Preparation: Invitrogen Corp

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

cDNA library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Place: LLNL4284 row: n column: 20
High quality sequence start: 31
High quality sequence stop: 446.
Location/Qualifiers

FEATURES

source

1. 959
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:672324"
/cell_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231,
hTERT-HMEL, INCAP"
/lab_host="BMDH10B"
/clone_lib="MAPCL"
/note="Vector: PCMV-SPORT6; Site 1: EcoRV; Site 2: Not I;
Subtracted with brain, liver, lung, kidney and muscle.
Directionally cloned. Priming method: oligo-dt. Average
insert size: 1800 bp. Library amplification: 26,000 fold.
Kristi A. Eglant, James V. Vincent, Robert Strausberg,
Bungkok Lee & Ira Pastan: Discovery of new breast
cancer genes encoding membrane and secreted proteins.
Manuscript submitted."

ORIGIN

Query Match 100.0%; Score 60; DB 6; Length 959;
Best Local Similarity 100.0%; Pred. No. 5.7e-06;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGCTCACCAGCCCCCAGCCAGCGTGTCACCTCGGCGGAGACACAGCGGCCCG 60
|||||
Db 349 GGCTCACCAGCCCCCAGCCAGCGTGTCACCTCGGCGGAGACACAGCGGCCCG 408

RESULT 5 BUI48487 1113 bp mRNA linear EST 03-SEP-2002
LOCUS AGENCOURT_8670479 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6380642
DEFINITION 5', mRNA sequence.
ACCESSION BUI48487
VERSION BUI48487.1 GI:22662019
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 1113)
NIH-MGC <http://mgi.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: DCTD/DRP
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Place: LNCM2569 row: j column: 03
High quality sequence stop: 235.
Location/Qualifiers

FEATURES

source

1. 1113
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6380642"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"

/note="Organ: prostate; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dt priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN

Query Match 100.0%; Score 60; DB 5; Length 1113;
Best Local Similarity 100.0%; Pred. No. 5.6e-06;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGCTCACCAGCCCCCAGCCAGCGTGTCACCTCGGCGGAGACACAGCGGCCCG 60
|||||
Db 90 GGCTCACCAGCCCCCAGCCAGCGTGTCACCTCGGCGGAGACACAGCGGCCCG 149

RESULT 6 BUI42996 1130 bp mRNA linear EST 13-SEP-2002
LOCUS AGENCOURT_10338707 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6574979
DEFINITION 5', mRNA sequence.
ACCESSION BUI42996
VERSION BUI42996.1 GI:22853479
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 1130)
NIH-MGC <http://mgi.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: DCTD/DRP
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Place: LNCM2771 row: c column: 11
High quality sequence stop: 27.
High quality sequence stop: 246.
Location/Qualifiers

FEATURES

source

1. 1130
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6574979"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dt priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN

Query Match 100.0%; Score 60; DB 5; Length 1130;
Best Local Similarity 100.0%; Pred. No. 5.6e-06;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGCTCACCAGCCCCCAGCCAGCGTGTCACCTCGGCGGAGACACAGCGGCCCG 60
|||||
Db 69 GGCTCACCAGCCCCCAGCCAGCGTGTCACCTCGGCGGAGACACAGCGGCCCG 128

RESULT 7
BQ936898 1234 bp mRNA linear EST 21-AUG-2002
LOCUS BQ936898
DEFINITION AGENCOURT_8919110 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6383855
5', mRNA sequence.
ACCESSION BQ936898
VERSION BQ936898.1 GI:22352281
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1234)
NIH-MGC <http://mgi.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Rubin Laboratory
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L1CM2577 row: 0 column: 24
High quality sequence stop: 245.
Location/Qualifiers
1..1234
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6383855"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_1ib="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC library."

ORIGIN
Query Match 100.0%; Score 60; DB 5; Length 1234;
Best Local Similarity 100.0%; Pred. No. 5.6e-06;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCTCACCGCCCCCGACGCGTGTCACTCGGCCCGACACGAGCGGCCCG 60
DB 90 GGCTCACCGCCCCCGACGCGTGTCACTCGGCCCGACACGAGCGGCCCG 149

RESULT 8
BQ943554 1268 bp mRNA linear EST 21-AUG-2002
LOCUS BQ943554
DEFINITION AGENCOURT_8777815 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6384308
5', mRNA sequence.
ACCESSION BQ943554
VERSION BQ943554.1 GI:22359032
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1268)
NIH-MGC <http://mgi.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Rubin Laboratory
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L1CM2579 row: b column: 21
High quality sequence stop: 177.
Location/Qualifiers
1..1268
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6384308"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_1ib="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC library."

ORIGIN
Query Match 100.0%; Score 60; DB 5; Length 1268;
Best Local Similarity 100.0%; Pred. No. 5.5e-06;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCTCACCGCCCCCGACGCGTGTCACTCGGCCCGACACGAGCGGCCCG 60
DB 90 GGCTCACCGCCCCCGACGCGTGTCACTCGGCCCGACACGAGCGGCCCG 149

RESULT 9
BQ920055 1343 bp mRNA linear EST 20-AUG-2002
LOCUS BQ920055
DEFINITION AGENCOURT_10031674 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6481609
5', mRNA sequence.
ACCESSION BQ920055
VERSION BQ920055.1 GI:22334753
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1343)
NIH-MGC <http://mgi.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Rubin Laboratory
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L1CM2663 row: e column: 02
High quality sequence start: 56
High quality sequence stop: 237.
Location/Qualifiers
1..1343
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6481609"

FEATURES
source

/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_1ib="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC library."

ORIGIN

Query Match 100.0%; Score 60; DB 5; Length 1343;
Best Local Similarity 100.0%; Pred. No. 5.5e-06;

Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GGCTCCACCGCCCCCGACCGGTGTACCTCGGCCCGGACACGAGCGGCCCG 60
132 GGCTCCACCGCCCCCGACCGGTGTACCTCGGCCCGGACACGAGCGGCCCG 191

RESULT 10
LOCUS BUI52566 1349 bp mRNA linear EST 03-SEP-2002
DEFINITION AGENCOURT_8670683 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6380559
ACCESSION BUI52566
VERSION BUI52566.1 GI:22666098
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 1349)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999).
CONTACT: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTF

CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LCM2569 row: f column: 16
High quality sequence stop: 291.
Location/Qualifiers

FEATURES

source

1. 1349
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6380559"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_1ib="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC library."

ORIGIN

Query Match 100.0%; Score 60; DB 5; Length 1349;
Best Local Similarity 100.0%; Pred. No. 5.5e-06;

Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GGCTCCACCGCCCCCGACCGGTGTACCTCGGCCCGGACACGAGCGGCCCG 60

19 GGCTCCACCGCCCCCGACCGGTGTACCTCGGCCCGGACACGAGCGGCCCG 78

RESULT 11
LOCUS BUI42790 1420 bp mRNA linear EST 13-SEP-2002
DEFINITION AGENCOURT_10327072 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6574725
ACCESSION BUI42790
VERSION BUI42790.1 GI:22853273
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 1420)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999).
CONTACT: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTF

CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LCM2770 row: h column: 21
High quality sequence stop: 288.
Location/Qualifiers

FEATURES

source

1. 1420
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6574725"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_1ib="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC library."

ORIGIN

Query Match 100.0%; Score 60; DB 5; Length 1420;
Best Local Similarity 100.0%; Pred. No. 5.5e-06;

Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GGCTCCACCGCCCCCGACCGGTGTACCTCGGCCCGGACACGAGCGGCCCG 60
19 GGCTCCACCGCCCCCGACCGGTGTACCTCGGCCCGGACACGAGCGGCCCG 78

RESULT 12
LOCUS BUI43309 1511 bp mRNA linear EST 13-SEP-2002
DEFINITION AGENCOURT_10327072 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6575349
ACCESSION BUI43309
VERSION BUI43309.1 GI:22853792
KEYWORDS EST.

SOURCE

Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 1511)
NIH-MGC http://mgi.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov

Tissue Procurement: DCTD/DTF
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM2772 row: b column: 21
High quality sequence stop: 166.

FEATURES

source
1. .1531
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6575349"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 40"
/note="Organ: prostate; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 97.3%; Score 58.4; DB 5; Length 1531;
Best Local Similarity 98.3%; Pred. No. 1.4e-05;
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGCTCCACCGCCCGCCAGCCAGCTGCTGCGCCCGGACACCGCGGCCCG 60
|||||
Db 30 GGCTCCACCGCCCGCCAGCCAGCTGCTGCGCCCGGACACCGCGGCCCG 89

RESULT 13
BQ935496 1262 bp mRNA linear EST 21-AUG-2002
LOCUS AGENCOURT_8776475 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6382088
DEFINITION 5', mRNA sequence.
ACCESSION BQ935496
VERSION BQ935496.1 GI:22350879
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1262)
NIH-MGC http://mgi.nci.nih.gov/.

REFERENCE NIH-MGC
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DTF
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM2573 row: f column: 09
High quality sequence start: 46
High quality sequence stop: 157.

FEATURES

source
1. .1262
/organism="Homo sapiens"
/mol_type="mRNA"

/db_xref="taxon:9606"
/clone="IMAGE:6382088"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 94.7%; Score 56.8; DB 5; Length 1262;
Best Local Similarity 96.7%; Pred. No. 3.8e-05;
Matches 58; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGCTCCACCGCCCGCCAGCCAGCTGCTGCGCCCGGACACCGCGGCCCG 60
|||||
Db 56 GGCTCCACCGCCCGCCAGCCAGCTGCTGCGCCCGGACACCGCGGCCCG 115

RESULT 14
BG774910 981 bp mRNA linear EST 15-MAY-2001
LOCUS 60264983221 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:4761054 5',
DEFINITION mRNA sequence.
ACCESSION BG774910
VERSION BG774910.1 GI:14045227
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 981)
NIH-MGC http://mgi.nci.nih.gov/.

REFERENCE NIH-MGC
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DTF
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM1612 row: o column: 07
High quality sequence stop: 874.

FEATURES

source
1. .981
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4761054"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 40"
/note="Organ: prostate; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 87.0%; Score 52.2; DB 4; Length 981;
Best Local Similarity 94.7%; Pred. No. 0.00062;
Matches 54; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGCTCACCAGCCCCCGGACCGGTCACCTCGAGCCCGGACACAGCCGCGCC 57
 Db 21 GGCTCACCAGCCCCCGGACCGGTCACCTCGAGCCCGGACACAGCCGCGCC 77

RESULT 15
 BQ923149 1536 bp mRNA 11near EST 20-AUG-2002
 LOCUS BQ923149
 DEFINITION AGENCOURT_8929207 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6484568
 5', mRNA sequence.
 ACCESSION BQ923149
 VERSION BQ923149.1 GI:22338180
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1536)
 NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-r@mail.nih.gov
 Tissue Procurement: DCTD/DRP
 CDNA Library Preparation: Rubin Laboratory
 DNA Sequencing By: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L1CM2670 row: P column: 09
 High quality sequence stop: 287.
 Location/Qualifiers
 1. 1536

FEATURES
 source
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6484568"
 /tissue_type="carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_1ib="NIH MGC 40"
 /note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
 Site_2: EcoRI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCAAGAG(G). Library constructed by
 King Hong in the laboratory of Gerald M. Rubin (University
 of California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC library."

ORIGIN

Query Match 87.0%; Score 52.2; DB 5; Length 1536;
 Best Local Similarity 94.7%; Pred. No. 0.00059;
 Matches 54; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 GGCTCACCAGCCCCCGGACCGGTCACCTCGAGCCCGGACACAGCCGCGCC 57
 Db 18 GGCTCACCAGCCCCCGGACCGGTCACCTCGAGCCCGGACACAGCCGCGCC 74

Search completed: January 16, 2005, 02:55:18
 uob time : 1555.6 secs

This Page Blank (uspto)